IRE Transactions

on INFORMATION THEORY

Volume IT-2 MARCH, 1956 Number 1

TABLE OF CONTENTS

PUBLISHED BY THE
Professional Group on Information Theory

IRE PROFESSIONAL GROUP ON INFORMATION THEORY

The Professional Group on Information Theory is an organization, within the framework of the IRE, of members with principal professional interest in Information Theory. All members of the IRE are eligible for membership in the Group and will receive all Group publications upon payment of the prescribed annual assessment of $2.00.

IRE TRANSACTIONS

on Information Theory

Published by the Institute of Radio Engineers, Inc., for the Professional Group on Information Theory at 1 East 79th Street, New York 21, N. Y. Responsibility for the contents rests upon the authors, and not upon the Institute, the Group or its members. Single copy prices: IRE-PGIT members, $1.60; IRE members, $2.40; nonmembers, $4.80.

Preliminary Announcement—Symposium on Information Theory

The Professional Group on Information Theory of the Institute of Radio Engineers, in cooperation with the Research Laboratory of Electronics of the Massachusetts Institute of Technology, is planning to hold a Symposium on Information Theory at the Massachusetts Institute of Technology, Cambridge, Mass., on September 10-12, 1956. The Symposium will be co-sponsored by the Office of Naval Research, the Air Research and Development Command, the Signal Corps Engineering Laboratories, and the U.R.S.I.

A similar symposium was conducted by the same organizations in September, 1954. It is the intention of the Organizing Committee to follow the pattern of the 1954 meeting, and to make this symposium an occasion at which authorities in this field may present and discuss the significant advances which have been made during the year. In order to provide an opportunity for informed and creative discussion of the papers, the committee is planning to distribute the Symposium Transactions two weeks prior to the meeting, as was done for the 1954 Symposium.

The Symposium Transactions will be distributed as a regular PGIT publication and copies will be available to the members of the co-sponsoring organizations at the same price as to all IRE members.

Submission of papers is hereby invited. In order to carry out successfully the advance publication plan, abstracts of the papers should be submitted at the earliest possible date. The form of the Symposium Transactions will be the same as that used for the IRE CONVENTION RECORD and authors will be expected to prepare a final copy of their papers on special format suitable for photo-offset reproduction. The material and instructions for the preparation of this final manuscript will be mailed to the author at the time of acceptance of the abstract or at any time upon request to the Chairman of the Organizing Committee: P. Elias, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge 39, Mass.

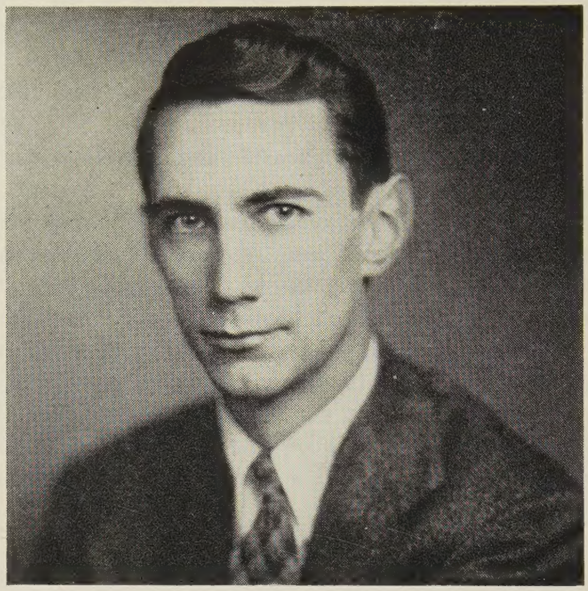
The deadline schedule for abstracts and full-length papers is as follows:

May 1—Deadline for receipt of 750-1,000 word abstract of proposed paper.
May 15—Authors notified of acceptance or rejection of papers submitted in abstract form.
June 15—Deadline for reception of full-length papers not previously submitted in abstract form. These must be submitted on special paper and in the approved format.
July 1—Deadline for reception of final master copy of papers which have been previously accepted in abstract form.
Sept. 1—Mailing of Symposium Transactions.

The above deadlines are rigid because of the publication schedule. Early submission of abstracts will be greatly appreciated by the Organizing Committee.

Annual ACM Meeting

The Annual Meeting of the Association for Computing Machinery will be held at the University of California Westwood Campus, Los Angeles, on August 27–29, 1956. For information, write to G. W. King, Box 3251, Olympic Station, Beverly Hills, Calif. Submit papers (abstract and four page manuscript in triplicate) by May 15 to J. P. Nash, University of Illinois, Urbana, Ill. (See January issue, *Journal of Association for Computing Machinery* for further details.)



CLAUDE E. SHANNON

Claude E. Shannon was born in Gaylord, Mich., on April 30, 1916. He received the B.S. degree in Electrical Engineering and Mathematics from the University of Michigan in 1936. From 1936 to 1940, he was at M.I.T., combining graduate studies with professional experience. For two years he was a research assistant in the Electrical Engineering Department, where he operated the Bush mechanical differential analyzer. He was an Assistant in the Mathematics Department from 1938 to 1940, and during 1939–1940, was a Bolles Fellow. He received the S.M. degree in Electrical Engineering and the Ph.D. degree in Mathematics from M.I.T. in 1940.

He was associated with the Institute for Advanced Study at Princeton University for one year thereafter as a National Research Fellow and consultant to the National Defense Research Committee. Since 1941, he has been a research mathematician for Bell Telephone Laboratories in Murray Hill, N. J.

Dr. Shannon's chief work has been in the following fields: the use of Boolean Algebra in relay and switching circuits, theory of communication, Mathematics of cryptography, theory of differential analyzer, and the use of computing machines for non-numerical operations. He also has studied chess-playing and maze-solving machines, the theory of Turing machines, design of reliable machines from unreliable components, stochastic processes, the Algebra of genetics, and graph theory.

In 1940, Dr. Shannon was the recipient of the Alfred Nobel Prize of the American Institute of Electrical Engineers for his work in switching theory. He received the Morris Liebmann award of the Institute of Radio Engineers in 1949 for his communication theory work. Yale University awarded him an honorary Master of Science degree in 1954, and in 1955, Dr. Shannon received the Stuart Ballantine medal of the Franklin Institute for work in communication theory. He is the author of approximately thirty-five technical papers, and holds several patents. He is co-author, with Warren Weaver, of "The Mathematical Theory of Communication," and co-editor, with John McCarthy, of the forthcoming "Automata Studies."

Dr. Shannon is a Fellow of the Institute of Radio Engineers, and a member of the American Mathematics Society, Sigma Xi, and Phi Kappa Phi.

The Bandwagon

CLAUDE E. SHANNON

INFORMATION theory has, in the last few years, become something of a scientific bandwagon. Starting as a technical tool for the communication engineer, it has received an extraordinary amount of publicity in the popular as well as the scientific press. In part, this has been due to connections with such fashionable fields as computing machines, cybernetics, and automation; and in part, to the novelty of its subject matter. As a consequence, it has perhaps been ballooned to an importance beyond its actual accomplishments. Our fellow scientists in many different fields, attracted by the fanfare and by the new avenues opened to scientific analysis, are using these ideas in their own problems. Applications are being made to biology, psychology, linguistics, fundamental physics, economics, the theory of organization, and many others. In short, information theory is currently partaking of a somewhat heady draught of general popularity.

Although this wave of popularity is certainly pleasant and exciting for those of us working in the field, it carries at the same time an element of danger. While we feel that information theory is indeed a valuable tool in providing fundamental insights into the nature of communication problems and will continue to grow in importance, it is certainly no panacea for the communication engineer or, *a fortiori*, for anyone else. Seldom do more than a few of nature's secrets give way at one time. It will be all too easy for our somewhat artificial prosperity to collapse overnight when it is realized that the use of a few exciting words like *information, entropy, redundancy,* do not solve all our problems.

What can be done to inject a note of moderation in this situation? In the first place, workers in other fields should realize that the basic results of the subject are aimed in a very specific direction, a direction that is not necessarily relevant to such fields as psychology, economics, and other social sciences. Indeed, the hard core of information theory is, essentially, a branch of mathematics, a strictly deductive system. A thorough understanding of the mathematical foundation and its communication application is surely a prerequisite to other applications. I personally believe that many of the concepts of information theory will prove useful in these other fields—and, indeed, some results are already quite promising—but the establishing of such applications is not a trivial matter of translating words to a new domain, but rather the slow tedious process of hypothesis and experimental verification. If, for example, the human being acts in some situations like an ideal decoder, this is an experimental and not a mathematical fact, and as such must be tested under a wide variety of experimental situations.

Secondly, we must keep our own house in first class order. The subject of information theory has certainly been sold, if not oversold. We should now turn our attention to the business of research and development at the highest scientific plane we can maintain. Research rather than exposition is the keynote, and our critical thresholds should be raised. Authors should submit only their best efforts, and these only after careful criticism by themselves and their colleagues. A few first rate research papers are preferable to a large number that are poorly conceived or half-finished. The latter are no credit to their writers and a waste of time to their readers. Only by maintaining a thoroughly scientific attitude can we achieve real progress in communication theory and consolidate our present position.

The Probability Distribution for the Filtered Output of a Multiplier Whose Inputs are Correlated, Stationary, Gaussian Time-Series

D. G. LAMPARD†

Summary—In this paper the techniques used by Kac and Siegert and by Emerson for evaluating the probability distribution for the filtered output of a square-law device with a stationary, Gaussian input, have been extended to the case of a multiplier whose inputs are a pair of correlated, stationary, Gaussian time-series. It is shown that in this case the probability distribution is determined by the eigenvalues of a pair of simultaneous, linear, homogeneous, integral equations whose kernels involve only the correlation functions of the inputs and the impulse response of the postmultiplier filter.

Explicit solutions for the eigenvalues of these integral equations are obtained both for the case of no postmultiplier filtering and for a simple example system using RC filters. Using these solutions the corresponding probability distributions are discussed and in particular, the way in which the probability distribution of the output tends to Gaussian as the postmultiplier filter time constant is increased, is demonstrated.

I. Introduction

A PAIR of stationary time series with Gaussian probability distributions are completely described statistically when their cross- and autocorrelation functions are specified.[1] Both for this reason and because of the relationship of the correlation functions to the corresponding power spectra,[2] the technique of measurement of such correlation functions has received a considerable amount of attention in the literature over the past few years.[3,4]

If we denote the time series by $f_1(t)$ and $f_2(t)$ the cross ($i \neq j$) and auto ($i = j$) correlation functions $\psi_{ij}(\tau)$ are defined by either

$$\Psi_{ij}(\tau) = \langle f_i(t) f_j(t - \tau) \rangle \tag{1}$$

$$= \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{+T} f_i(t) f_j(t - \tau)\, dt, \tag{2}$$

where in (1) the angular brackets denote ensemble averaging. The equivalence of these definitions follows from the ergodic hypothesis.[5]

When the $f_1(t)$, $f_2(t)$ are voltages or currents, a analog system of the type shown in Fig. 1 is commonl used to measure the correlation function $\psi_{12}(\tau)$.

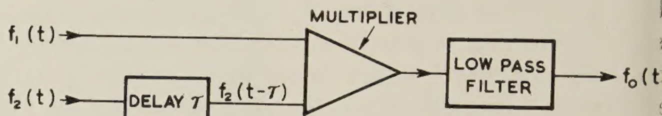


Fig. 1—Typical analog correlator.

In this system a low pass filter of finite time constan is used to carry out approximately, the time averagin operation implied in (2). It will be clear that the outpu $f_0(t)$ of such a system is a random function of tim whose statistical properties are a function of the tim constant of the low-pass filter (that is, for a given typ of low-pass filter). Thus the output exhibits fluctuatior about its mean $\psi_{12}(\tau)$, the magnitude of which decrease as the time constant of the low-pass filter is increased While a direct determination of the variance of thes fluctuations is fairly straightforward,[6] the determinatio of their probability distribution is a much more difficu problem. This difficulty arises because of the non-Gaussia distribution of the input to the low pass filter. (Se Section III).

From a statistical point of view the essential element of the system of Fig. 1 are the multiplier and the filte (not necessarily low-pass) and for this reason we sha concern ourselves with the system of Fig. 2.

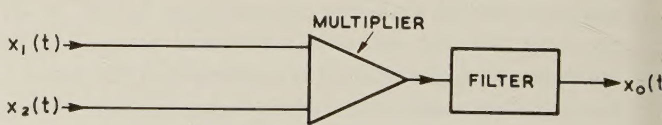


Fig. 2—Basic system for analysis.

For the special case in which the two input time seri are identical, $x_1(t) = x_2(t)$, this system has been studie by several authors, in particular Franz,[7] Kac and Sieger

† Commonwealth Scientific and Industrial Research Organization, Division of Electrotechnology, University Grounds, Chippendale, Sydney, Australia.

[1] Zero means are assumed without loss of generality.

[2] This relationship is generally known as the Wiener-Khintchine Theorem.

[3] Fairly complete bibliographies of correlation techniques will be found in F. L. Stumpers, "A Bibliography of Information Theory," M.I.T. Res. Lab. Elec. Tech. Rep., pp. 11–15; February, 1953 and supplement to "A Bibliography of Information Theory," N. V. Philips Gloeilampenfabrieken, Eindhoven, Netherlands, pp. 8–11; April, 1954.

[4] D. G. Lampard, "A new method of determining correlation functions of stationary time series," *Proc. IEE*, Part IV, Monograph No. 104, 1954.

[5] N. Wiener, "Extrapolation, Interpolation and Smoothing of Stationary Time Series," John Wiley & Sons, New York, p. 15; 1950.

[6] Assuming that the $f_1(t)$ and $f_2(t)$ are Gaussian. This assumptic is made throughout this paper.

[7] K. Franz, "Die amplituden von Geranchshanungen" *Elek. Nac Tech.*, vol. 19, p. 166; September, 1942. The author is grateful Mr. R. E. Burgess for bringing this reference to his attention.

[8] M. Kac and A. J. F. Siegert, "On the theory of noise in rad receivers with square law detectors," *Jour. Appl. Phys.*, vol. 1 pp. 383–397; April, 1947.

nd Emerson[9]. These authors have shown that the robability distribution of the output $x_0(t)$, is determined y the eigenvalues of a linear, homogeneous, integral quation whose kernel involves only the autocorrelation nction of the input and the impulse response of the ter. More recently Siegert[10] has pointed out that it is ten more convenient to work with a closely related homogeneous integral equation. In particular he has own that this inhomogeneous integral equation reduces a differential equation when the input has been obtained y filtering white noise with a finite lumped circuit etwork. The way in which the probability distribution ehaves in the limiting case of a very narrow band filter as been discussed by Arthur[11] while Meyer and Middleon[12] have shown how the treatment of the authors entioned above can be extended to obtain multimensional probability distributions.

In this paper we shall employ similar methods to study e general system of Fig. 2. It will be shown that in this ase a pair of simultaneous, linear, homogeneous, integral quations arise whose eigenvalues determine the probbility distribution of $x_0(t)$. The special case of no postultiplier filter will be discussed and finally we shall eat in detail a simple example system and show how xplicit solutions for the eigenfunctions and eigenvalues the integral equations can be obtained. By making use these solutions we demonstrate the way in which the utput probability distribution tends to Gaussian as the ostmultiplier filter time constant is increased.

II. Analysis

The filtered output of the multiplier may be written s a convolution integral,[13] thus:

$$x_0(t) = \int_0^t x_1(t-u)x_2(t-u)h(u)\,du. \tag{3}$$

Here the inputs $x_1(t)$, $x_2(t)$ are assumed to be correted stationary time series with Gaussian probability istributions while $h(t)$ is the impulse response of the postultiplier filter.

It will be convenient to introduce new variables defined y

$$\begin{aligned} X_1(t) &= \tfrac{1}{2}[x_1(t) + x_2(t)] \\ X_2(t) &= \tfrac{1}{2}[x_1(t) - x_2(t)]. \end{aligned} \tag{4}$$

[9] R. C. Emerson, "First probability densities for receivers with uare law detectors," *Jour. Appl. Phys.*, vol. 24, pp. 1168–1176; ptember, 1953.

[10] A. J. F. Siegert, "Passage of stationary processes through near and non-linear devices," Trans. IRE, vol. IT-3, pp. 4–25; arch, 1954.

[11] G. R. Arthur, "A note on the approach of narrow band noise ter a non-linear device to a normal probability density," *Jour. ppl. Phys.*, vol. 23, pp. 1143–1144; October, 1952.

[12] M. A. Meyer and D. Middleton, "On the distribution of signals d noise after rectification and filtering," *Jour. Appl. Phys.*, vol. 25, . 1037–1052; August, 1954.

[13] H. S. Carslaw and J. C. Jaeger, "Operational methods in applied athematics," 2nd ed., Oxford University Press, New York, p. 252; 48.

As these new variables are simply linear combinations of the $x_1(t)$ and $x_2(t)$, it follows, from a well-known theorem[14] that they also have stationary Gaussian probability distributions.

If further we denote their correlation functions by

$$\phi_{ij}(\tau) = \langle X_i(t)X_j(t+\tau)\rangle, \tag{5}$$

then it is easy to show that

$$\phi_{11}(\tau) = \tfrac{1}{4}\{\psi_{11}(\tau) + \psi_{12}(\tau) + \psi_{21}(\tau) + \psi_{22}(\tau)\} \tag{6}$$

$$\phi_{12}(\tau) = \tfrac{1}{4}\{\psi_{11}(\tau) - \psi_{12}(\tau) + \psi_{21}(\tau) - \psi_{22}(\tau)\} \tag{7}$$

$$\phi_{21}(\tau) = \tfrac{1}{4}\{\psi_{11}(\tau) + \psi_{12}(\tau) - \psi_{21}(\tau) - \psi_{22}(\tau)\} \tag{8}$$

$$\phi_{22}(\tau) = \tfrac{1}{4}\{\psi_{11}(\tau) - \psi_{12}(\tau) - \psi_{21}(\tau) + \psi_{22}(\tau)\} \tag{9}$$

where $\psi_{ij}(\tau) = \langle x_i(t)\, x_j(t+\tau)\rangle$ are the correlation functions for the original variables $x_1(t)$, $x_2(t)$.

Then in terms of the new variables $X_1(t)$, $X_2(t)$ we may write (3) as

$$x_0(t) = \int_0^t \{X_1^2(t-u) - X_2^2(t-u)\}h(u)\,du. \tag{10}$$

We shall in fact study the more general form

$$x_0(t) = \int_0^t \{X_1^2(t-u) \pm X_2^2(t-u)\}h(u)\,du \tag{11}$$

$$= \int_0^t X_1^2(t-u)h(u)\,du \pm \int_{-t}^0 X_2^2(t+u)h(-u)\,du. \tag{12}$$

To make further progress we introduce the discontinuous functions,

$$U_1(x) = \begin{cases} 1, & x > 0+ \\ 0, & x < 0+ \end{cases} \tag{13}$$

$$U_2(x) = \begin{cases} 0, & x > 0- \\ 1, & x < 0-, \end{cases} \tag{14}$$

and define new variables,

$$Z(u) = X_1(t-u)U_1(u) + X_2(t+u)U_2(u) \tag{15}$$

$$H(u) = h(u)U_1(u) \pm h(-u)U_2(u) \tag{16}$$

$$= h(|u|)\{U_1(u) \pm U_2(u)\}. \tag{17}$$

It is clear that

$$Z^2(u) = X_1^2(t-u)U_1(u) + X_2^2(t-u)U_2(u), \tag{18}$$

and so (12) can now be written

$$x_0(t) = \int_{-t}^{+t} Z^2(u)H(u)\,du \tag{19}$$

$$= \lim_{n\to\infty} \sum_{i=-n}^{+n} \sum_{j=-n}^{+n} B_{ij}Z_iZ_j \tag{20}$$

[14] S. S. Wilks, "Mathematical Statistics," Princeton University Press, p. 70; 1943.

where

$$Z_j \equiv Z(u_j) = Z\left(j\frac{t}{n}\right) \tag{21}$$

$$B_{ij} = \frac{t}{n}H\left(j\frac{t}{n}\right)\delta_{ij}. \tag{22}$$

Here δ_{ij} is the Kronecker delta and the dash on the summation indicates that the term $j = 0$ is to be excluded.

As $X_1(t)$, $X_2(t)$ have stationary Gaussian probability distributions, it follows that Z_j has a $2n$ dimensional Gaussian distribution. Thus we may write [15]

$$p(Z_{-n} \cdots Z_{-1}, Z_{+1} \cdots Z_{+n})$$

$$= (2\pi)^{-n} \mid A \mid^{1/2} \exp - \tfrac{1}{2} \sum_{-n}^{+n} \sum_{-n}^{+n} A_{ij}Z_iZ_j, \tag{23}$$

where $[A]$ is the inverse of the second moment matrix $[M]$ whose typical element M_{ij} is given by

$$M_{ij} = \langle Z_iZ_j \rangle = \langle Z(u_i)Z(u_j) \rangle. \tag{24}$$

Now let us find the characteristic function $M(i\theta)$ for $x_0(t)$ defined by

$$M(i\theta) = \langle e^{i\theta x_0(t)} \rangle \tag{25}$$

$$= \lim_{n\to\infty} (2\pi)^{-n} \mid A \mid^{1/2}$$

$$\cdot \overbrace{\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty}}^{2n\text{fold}} \exp \{ -\tfrac{1}{2} \sum \sum A_{ij}Z_iZ_j$$

$$+ i\theta \sum \sum B_{ij}Z_iZ_j \} \, dZ_{-n} \cdots dZ_n. \tag{26}$$

As $[A]$ is the coefficient matrix for a Gaussian Multivariate distribution it is positive definite and symmetric. $[B]$ is not necessarily definite but these conditions are sufficient for a simultaneous reduction of the quadratic forms, occurring in (26), to principal axes.[16]

Thus there exists a linear transformation

$$Z_i = \sum_{j=-n}^{+n} {}_{ij}y_j \tag{27}$$

such that

$$\sum_{i=-n,}^{+n} \sum_{j=-n}^{+n} A_{ij}Z_iZ_j = \sum_{j=-n}^{+n} y_j^2 \tag{28}$$

$$\sum_{i=-n}^{+n} \sum_{i=-n}^{+n} B_{ij}Z_iZ_j = \sum_{j=-n}^{+n} \lambda_j y_j^2, \tag{29}$$

where λ_j are roots of the equation

$$\mid B_{ij} - \lambda A_{ij} \mid = 0. \tag{30}$$

The Jacobian of the transformation (27) is just

$$J = \left| \frac{\partial(Z_{-n}, \cdots Z_n)}{\partial(y_{-n}, \cdots y_n)} \right| = \mid l_{ij} \mid, \tag{31}$$

[15] S. S. Wilks, *loc. cit.*, p. 63.

[16] R. Courant and D. Hilbert, "Methods of Mathematical Physics," Vol. 1, Interscience Publishers, New York, p. 37; 1953.

so that (26) may be written

$$M(i\theta) = \lim_{n\to\infty} \mid A \mid^{1/2} \mid l \mid \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} (2\pi)^{-n}$$

$$\cdot \exp \{ -\tfrac{1}{2} \sum (1 - 2i\theta\lambda_j)y_j^2 \} \, dy_{-n} \cdots dy_n \quad (3$$

$$= \lim_{n\to\infty} \mid A \mid^{1/2} \mid l \mid \prod_{j=-n}^{+n} \frac{1}{\sqrt{2\pi}}$$

$$\int_{-\infty}^{+\infty} \exp \{ -\tfrac{1}{2} (1 - 2i\theta\lambda_j)y_j^2 \} \, dy_j \quad (3$$

$$= \lim_{n\to\infty} \mid A \mid^{1/2} \mid l \mid \prod_{j=-n}^{+n} (1 - 2i\theta\lambda_j)^{-1/2}. \quad (3$$

However it is clear from (25) that $M(0) = 1$ so that w must have $\mid A \mid^{\frac{1}{2}} \mid l \mid = 1$ and hence finally we may wri

$$M(i\theta) = \lim_{n\to\infty} \prod_{j=-n}^{+n} (1 - 2i\theta\lambda_j)^{-1/2}. \quad (3$$

We now find the cumulants[17] (or semi-invariants) of tl probability distribution of $x_0(t)$. The mth cumulant defined as the coefficient of $(i\theta)^m/m!$ in the expansion the logarithim of the characteristic function. Thus w have

$$\log M(i\theta) = \log \left\{ \lim_{n\to\infty} \prod_{j=-n}^{+n} (1 - 2i\theta\lambda_j)^{-1/2} \right\} \quad (3$$

$$= \lim_{n\to\infty} \left\{ \frac{1}{2} \sum_{m=1}^{\infty} \frac{(i\theta)^m}{m} \left[2^m \sum_{j=-n}^{+n} \lambda_j^m \right] \right\} \quad (3$$

$$= \sum_{m=1}^{\infty} \frac{(i\theta)^m}{m!} K_m, \quad (3$$

where

$$K_m = 2^{m-1}(m-1)! \left\{ \lim_{n\to\infty} \sum_{j=-n}^{+n} \lambda_j^m \right\}. \quad (3$$

Eq. (39) shows that the cumulants of the probabili distribution for $x_0(t)$ are simply proportional to the su of the mth powers of the roots of (30). When we use (2 and (22) we see that (30) may be written

$$\left| \frac{t}{n} H(u_j)\delta_{ij} - \lambda A_{ij} \right| = 0; \quad (4$$

that is

$$\left| \lambda\delta_{ij} - \frac{t}{n} H(u_i)M_{ij} \right| = 0, \quad (4$$

where, as before, M_{ij} is an element of the second mome matrix defined by (24).

In the limit, as $n \to \infty$, putting $u_i = u$, $u_j = v$, (2 becomes

$$M_{ij} \to \langle \{X_1(t-u)U_1(u) + X_2(t+u)U_2(u)\}$$

$$\cdot \{X_1(t-v)U_1(v) + X_2(t-v)U_2(v)\} \rangle \quad (4$$

[17] H. Cramér, "Mathematical Methods of Statistics," Princet University Press, p. 186; 1951.

$$= \phi_{11}(u - v)U_1(u)U_1(v) + \phi_{12}(u + v)U_1(u)U_2(v) + \phi_{21}(-v - u)U_2(u)U_1(v) + \phi_{22}(v - u)U_2(u)U_2(v) \quad (43)$$

$$= \Phi(u, v) \quad (44)$$

From the theory of integral equations,[18] we see at once hat (41) is just the Fredholm determinant of the linear iomogeneous integral equation

$$\lambda f(v) = \int_{-t}^{+t} \Phi(u, v)H(u)f(u)\, du. \quad (45)$$

In particular, when a sufficiently long time has elapsed or a statistically steady state to be reached, we may xtend the limits to infinity obtaining the singular ntegral equation

$$\lambda f(v) = \int_{-\infty}^{+\infty} \Phi(u, v)H(u)f(u)\, du \quad (46)$$

$$= \int_{-\infty}^{+\infty} K(u, v)f(u)\, du, \quad (47)$$

vhere

$$K(u, v) = h(|u|)\{\phi_{11}(u - v)U_1(u)U_1(v) + \phi_{12}(u + v)U_1(u)U_2(v) \pm \phi_{21}(-v - u)U_2(u)U_1(v) \pm \phi_{22}(v - u)U_2(u)U_2(v)\}. \quad (48)$$

f now we set

$$f(v) = f_1(v)U_1(v) + f_2(-v)U_2(v), \quad (49)$$

ve find that (48) may be written as the pair of simultaneous integral equations

$$\begin{aligned} \lambda f_1(v) &= \int_0^\infty \phi_{11}(v - u)h(u)f_1(u)\, du \pm \int_0^\infty \phi_{12}(v - u)h(u)f_2(u)\, du. \\ \lambda f_2(v) &= \int_0^\infty \phi_{21}(v - u)h(u)f_1(u)\, du \pm \int_0^\infty \phi_{22}(v - u)h(u)f_2(u)\, du. \end{aligned} \quad (50)$$

inally we write

$$\begin{aligned} f_1(v) &= \tfrac{1}{2}\{g_1(v) + g_2(v)\} \\ f_2(v) &= \tfrac{1}{2}\{g_1(v) - g_2(v)\} \end{aligned} \quad (51)$$

nd consider the plus and minus cases separately. We ind on making use of (6), (7), (8) and (9), that (50) may e written as follows:

Case 1 $+^{ve}$ Filter Input $= \frac{1}{2}[x_1^2(t) + x_2^2(t)]$

$$\begin{aligned} 2\lambda g_1(v) &= \int_0^\infty \psi_{11}(v - u)h(u)g_1(u)\, du + \int_0^\infty \psi_{12}(v - u)h(u)g_2(u)\, du \\ 2\lambda g_2(v) &= \int_0^\infty \psi_{21}(v - u)h(u)g_1(u)\, du + \int_0^\infty \psi_{22}(v - u)h(u)g_2(u)\, du. \end{aligned} \quad (52)$$

Case 2 $-^{ve}$ Filter Input $= x_1(t)\, x_2(t)$

$$\begin{aligned} 2\lambda g_1(v) &= \int_0^\infty \psi_{12}(v - u)h(u)g_1(u)\, du + \int_0^\infty \psi_{11}(v - u)h(u)g_2(u)\, du \\ 2\lambda g_2(v) &= \int_0^\infty \psi_{22}(v - u)h(u)g_1(u)\, du + \int_0^\infty \psi_{21}(v - u)h(u)g_2(u)\, du. \end{aligned} \quad (53)$$

Eq. (53) involves only the correlation functions of the inputs to the multiplier and the impulse response of the postmultiplier filter.

In particular we note that, if we set $x_1(t) = x_2(t)$, (53) reduces immediately to the single integral equation

$$\lambda g(v) = \int_0^\infty \psi(v - u)h(u)g(u)\, du, \quad (54)$$

which is identical with that given by Kac and Siegert[8] and Franz[7] for the square law detector (after appropriate change of notation).

III. The Case of No Postmultiplier Filter

In this section we shall make use of (53) to evaluate the probability distribution for the output $x_0(t)$ in the special case of no postmultiplier filtering. In this case the impulse response $h(t)$ of the filter is just replaced by the impulse function $\delta(t)$ and the integrations carried out. We find

$$\begin{aligned} 2\lambda g_1(v) &= \psi_{12}(v)g_1(0) + \psi_{11}(v)g_2(0) \\ 2\lambda g_2(v) &= \psi_{22}(v)g_1(0) + \psi_{21}(v)g_2(0). \end{aligned} \quad (55)$$

Setting $v = 0$ and noting that $\psi_{12}(0) = \psi_{21}(0)$, we may write (55) in the form

$$\begin{aligned} [2\lambda - \psi_{12}(0)]g_1(0) &= \psi_{11}(0)g_2(0), \\ [2\lambda - \psi_{12}(0)]g_2(0) &= \psi_{22}(0)g_1(0), \end{aligned} \quad (56)$$

and hence we must have

$$[2\lambda - \psi_{12}(0)]^2 = \psi_{11}(0)\psi_{22}(0), \quad (57)$$

[18] W. V. Lovitt, "Linear Integral Equations," Dover Publications, Vew York, p. 24; 1950.

which is the eigenvalue equation whose solutions are

$$\lambda = \frac{\psi_{12}(0) \pm \sqrt{\psi_{11}(0)\psi_{22}(0)}}{2}. \tag{58}$$

Let us write

$$\begin{aligned} \lambda_1' &= \frac{\psi_{12}(0) + \sqrt{\psi_{11}(0)\psi_{22}(0)}}{2} \\ \lambda_2' &= \frac{\sqrt{\psi_{11}(0)\psi_{22}(0)} - \psi_{12}(0)}{2}. \end{aligned} \tag{59}$$

Then it follows from the inequality[19]

$$\psi_{11}(0)\psi_{22}(0) - \psi_{12}^2(0) \geq 0 \tag{60}$$

that both λ_1', λ_2' are positive. Then we see from (35) that the characteristic function has the simple form

$$M(i\theta) = \{(1 - 2i\theta\lambda_1')(1 + 2i\theta\lambda_2')\}^{-1/2}, \tag{61}$$

and the corresponding probability density is given by the Fourier transform:

$$p(x_0) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} d\theta\, e^{-ix_0\theta}\{(1 - 2i\theta\lambda_1')(1 + 2i\theta\lambda_2')\}^{-1/2}. \tag{62}$$

Fortunately this integral is readily evaluated in closed form.

To carry out this evaluation we first complete the square in the denominator of (62) obtaining

$$p(x_0) = \frac{1}{4\pi}(\lambda_1'\lambda_2')^{-1/2}\int_{-\infty}^{+\infty} d\theta\, e^{-ix_0\theta}\left\{\left(\frac{\lambda_1' + \lambda_2'}{4\lambda_1'\lambda_2'}\right)^2 + \left[\theta + \frac{i(\lambda_2' - \lambda_1')}{4\lambda_1'\lambda_2'}\right]^2\right\}^{-1/2}. \tag{63}$$

On changing the variable by the substitution

$$\theta + \frac{i(\lambda_2' - \lambda_1')}{4\lambda_1'\lambda_2'} = z,$$

followed by shifting the contour of integration back to the real axis (this step is easily justified) we find

$$p(x_0) = \frac{1}{4\pi}(\lambda_1'\lambda_2')^{-1/2}\exp\left\{-x_0\frac{(\lambda_2' - \lambda_1')}{4\lambda_1'\lambda_2'}\right\} \cdot \int_{-\infty}^{+\infty} dz\, e^{-ix_0z}\left\{\left(\frac{\lambda_1' + \lambda_2'}{4\lambda_1'\lambda_2'}\right)^2 + z^2\right\}^{-1/2}. \tag{64}$$

Now, Bassett's integral representation[20] for the modified Bessel function of the second kind and zero order is

$$K_0(a \mid x \mid) = \frac{1}{2}\int_{-\infty}^{+\infty} \frac{e^{-ixz}}{(a^2 + z^2)^{1/2}}\, dz, \tag{65}$$

so using (65) in (64), we have

$$p(x_0) = \frac{1}{2\pi}(\lambda_1'\lambda_2')^{-1/2}\exp\left\{-x_0\frac{(\lambda_2' - \lambda_1')}{4\lambda_1'\lambda_2'}\right\} \cdot K_0\left\{\mid x_0 \mid \frac{\lambda_1' + \lambda_2'}{4\lambda_1'\lambda_2'}\right\}. \tag{66}$$

We now substitute the values given by (59) for λ_1', λ and obtain finally

$$p(x_0) = \frac{1}{\pi}\{\psi_{11}(0)\psi_{22}(0) - \psi_{12}^2(0)\}^{-1/2} \cdot \exp\left\{x_0\frac{\psi_{12}(0)}{\psi_{11}(0)\psi_{22}(0) - \psi_{12}^2(0)}\right\} \cdot K_0\left\{\mid x_0 \mid \frac{\sqrt{\psi_{11}(0)\psi_{22}(0)}}{\psi_{11}(0)\psi_{22}(0) - \psi_{12}^2(0)}\right\}. \tag{67}$$

This result, which gives the one-dimensional probabilit distribution for the product of two correlated Gaussia random variables, is a special case ($n = 1$) of a slightl more general result given by Wishart and Bartlett,[21] an is well known in classical statistics. We note also, that in recent interesting paper (which has several points c contact with the present work) G. R. Arthur[22] discusse the integral which occurs in (62) and shows that it may b evaluated in terms of Whittaker's confluent hyper geometric function. When it is realized, however, that th particular hypergeometric function encountered is readil expressible in terms of a Bessel function, some simpli fication results.

IV. Explicit Solution of a Simple Example System

In this section we shall consider in detail the followin system (see Fig. 3).

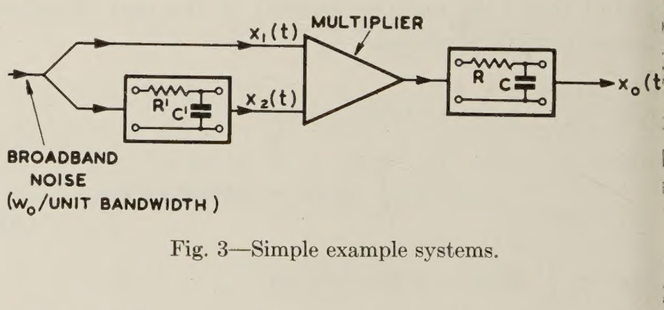


Fig. 3—Simple example systems.

The impulse response of the postmultiplier filter is jus

$$h(t) = \beta e^{-\beta t}, \qquad t > 0 \quad \text{where} \quad \beta = \frac{1}{RC}, \tag{68}$$

while it is easy to show that the correlation function $\psi_{ij}(\tau)$ are given by

[19] This follows immediately from the Schwarz inequality applied to the temporal definitions of the correlation functions.

[20] G. N. Watson, "Theory of Bessel Functions," Cambridge University Press, p. 172; 1948.

[21] J. S. Wishart and M. S. Bartlett, "The distribution of secon order moment statistics in a normal system," *Proc. Cambridge Phi Soc.*, vol. 28, p. 455; 1932.

[22] G. R. Arthur, "The statistical properties of the output of frequency sensitive device," *Jour. Appl. Phys.*, vol. 25, p. 118 September, 1954.

$$\psi_{11}(\tau) = \frac{\omega_0}{2}\,\delta(\tau), \qquad \text{all } \tau \tag{69}$$

$$\begin{aligned}\psi_{12}(\tau) &= \frac{\omega_0}{2}\,\alpha e^{-\alpha\tau}, && \tau > 0\\ &= 0 && \tau < 0\end{aligned} \tag{70}$$

$$\begin{aligned}\psi_{21}(\tau) &= 0, && \tau > 0\\ &= \frac{\omega_0}{2}\,\alpha e^{\alpha\tau}, && \tau < 0\end{aligned} \tag{71}$$

$$\psi_{22}(\tau) = \frac{\omega_0}{2}\cdot\frac{\alpha}{2}\,e^{-\alpha|\tau|}, \qquad \text{all } \tau, \tag{72}$$

where $\delta(\tau)$ in (69) is the even impulse function and where $\alpha = 1/R'C'$.

On making use of (68) through (72), in (53) we obtain

$$\mu g_1(v) = \alpha\beta e^{-\alpha v}\int_0^v e^{(\alpha-\beta)u}g_1(u)\,du + \beta e^{-\beta v}g_2(v) \tag{73}$$

$$\begin{aligned}\mu g_2(v) = \frac{\alpha\beta}{2}\,e^{-\alpha v}\int_0^v e^{(\alpha-\beta)u}g_1(u)\,du\\ + \frac{\alpha\beta}{2}\,e^{\alpha v}\int_v^\infty e^{-(\alpha+\beta)u}g_1(u)\,du\\ + \alpha\beta e^{\alpha v}\int_v^\infty e^{-(\alpha+\beta)u}g_2(u)\,du,\end{aligned} \tag{74}$$

where we have written

$$\mu = \frac{4\lambda}{\omega_0} \tag{75}$$

We now solve the simultaneous integral equations (73) and (74) by reduction to differential equations. This technique is essentially that employed by Juncosa[23] to solve the single integral equation which arose in an example in the work of Kac and Siegert. It should be pointed out, however, that the solutions were first obtained by the author as a series expansion[24] in powers of $e^{-\beta v}$. The series expansion was recognized as that of a Bessel function and hence it was realized that a more direct approach would be obtained by showing that the solutions satisfied Bessel's differential equation.

We differentiate (74) with respect to v and obtain

$$\begin{aligned}\mu g_2'(v) = -\frac{\alpha^2\beta}{2}\,e^{-\alpha v}\int_0^v e^{(\alpha-\beta)u}g_1(u)\,du\\ + \frac{\alpha^2\beta}{2}\,e^{\alpha v}\int_v^\infty e^{-(\alpha+\beta)u}g_1(u)\,du\\ + \alpha^2\beta e^{\alpha v}\int_v^\infty e^{-(\alpha+\beta)u}g_2(u)\,du - \alpha\beta e^{-\beta v}g_2(v).\end{aligned} \tag{76}$$

Multiplying (74) by $-\alpha$ and adding to (76) gives

$$\mu[g_2'(v) - \alpha g_2(v)]$$
$$= -\alpha^2\beta e^{-\alpha v}\int_0^v e^{(\alpha-\beta)u}g_1(u)\,du - \alpha\beta e^{-\beta v}g_2(v) \tag{77}$$
$$= -\alpha\{\mu g_1(v) - \beta e^{-\beta v}g_2(v)\} - \alpha\beta e^{-\beta v}g_2(v) \tag{78}$$
$$= -\alpha\mu g_1(v). \tag{79}$$

In passing from (77) to (78) we have made use of (73). Thus we have

$$g_2'(v) - \alpha g_2(v) = -\alpha g_1(v). \tag{80}$$

Again, differentiating (77) with respect to v gives

$$\mu[g_2''(v) - \alpha g_2'(v)]$$
$$\begin{aligned}= \alpha^3\beta e^{-\alpha v}\int_0^v e^{(\alpha-\beta)u}g_1(u)\,du - \alpha^2\beta e^{-\beta v}g_1(v)\\ - \alpha\beta e^{-\beta v}[g_2'(v) - \beta g_2(v)]\end{aligned} \tag{81}$$
$$\begin{aligned}= \alpha^2[\mu g_1(v) - \beta e^{-\beta v}g_2(v)] - \alpha^2\beta e^{-\beta v}g_1(v)\\ - \alpha\beta e^{-\beta v}[g_2'(v) - \beta g_2(v)],\end{aligned} \tag{82}$$

where again use has been made of (73) to remove the integral on the right-hand side of (81).

Now multiply (80) by $\alpha[\mu - \beta e^{-\beta v}]$ and add to (82). We find on collecting terms and dividing by μ,

$$g_2''(v) + \left\{\frac{\alpha\beta}{\mu}(2\alpha - \beta)e^{-\beta v} - \alpha^2\right\}g_2(v) = 0. \tag{83}$$

On making the substitutions,

$$x = 2\sqrt{\frac{\alpha}{\mu}\left(\frac{2\alpha}{\beta} - 1\right)}\cdot\exp\left\{-\frac{\beta}{2}v\right\} \tag{84}$$

$$g_2(v) = f\left\{2\sqrt{\frac{\alpha}{\mu}\left(\frac{2\alpha}{\beta} - 1\right)}\exp\left\{-\frac{\beta}{2}v\right\}\right\}, \tag{85}$$

it is easy to show that (83) may be written

$$x^2 f''(x) + x f'(x) + \left\{x^2 - \left(\frac{2\alpha}{\beta}\right)^2\right\}f(x) = 0; \tag{86}$$

which latter equation will be recognized as Bessel's differential equation of order $2\alpha/\beta$, whose solution is

$$\begin{aligned}f(x) &= AJ_{2\alpha/\beta}(x) + BJ_{-2\alpha/\beta}(x), && \frac{2\alpha}{\beta} \neq \text{an integer}\\ &= AJ_{2\alpha/\beta}(x) + BY_{2\alpha/\beta}(x), && \frac{2\alpha}{\beta} = \text{an integer}\end{aligned} \tag{87}$$

where A, B are arbitrary constants.

It will be seen from (74) however, that $g_2(v) \to 0$ as $v \to \infty$ and hence it is clear that we must have $B = 0$. Thus finally we may write

$$g_2(v) = AJ_{2\alpha/\beta}\left\{2\sqrt{\frac{\alpha}{\mu}\left(\frac{2\alpha}{\beta} - 1\right)}\cdot\exp\left\{-\frac{\beta}{2}v\right\}\right\}, \tag{88}$$

[23] M. L. Juncosa, "An integral equation related to Bessel functions," *Duke Math. Jour.*, vol. 12, p. 465; September, 1945.

[24] Quite recently an interesting note has appeared on the series solution of integral equations. See J. R. M. Radok, "The solution of eigenvalue problems of integral equations by power series," *Quart. Appl. Math.*, vol. 12, p. 413; January, 1955.

with now no restrictions as to whether $2\alpha/\beta$ is an integer or not. To find $g_1(v)$ we use (88) in (80). Performing the differentiation and making use of the recurrence relation

$$xJ_\nu'(x) + \nu J_\nu(x) = xJ_{\nu-1}(x), \tag{89}$$

we find readily that

$$g_1(v) = A\cdot\frac{\beta}{\alpha}\cdot\sqrt{\frac{\alpha}{\mu}\left(\frac{2\alpha}{\beta}-1\right)}\exp\left\{-\frac{\beta}{2}v\right\}$$
$$\cdot J_{(2\alpha/\beta)-1}\left\{2\sqrt{\frac{\alpha}{\mu}\left(\frac{2\alpha}{\beta}-1\right)}\exp\left\{-\frac{\beta}{2}v\right\}\right\}. \tag{90}$$

To determine the eigenvalues we evaluate the expression

$$I = \alpha\beta e^{-\alpha v}\int_0^v e^{(\alpha-\beta)u}g_1(u)\,du + \beta e^{-\beta v}g_2(v), \tag{91}$$

which is just the right-hand side of (73). To do this, we use the series expansions for the Bessel functions occurring in (88) and (90), and integrate termwise in (91).

We find, after some reduction,

$$I = \mu A\cdot\frac{\beta}{\alpha}\sqrt{\frac{\alpha}{\mu}\left(\frac{2\alpha}{\beta}-1\right)}\exp\left\{-\frac{\beta}{2}v\right\}$$
$$\cdot J_{(2\alpha/\beta)-1}\left\{2\sqrt{\frac{\alpha}{\mu}\left(\frac{2\alpha}{\beta}-1\right)}\exp\left\{-\frac{\beta}{2}v\right\}\right\}.$$
$$-\ \beta\cdot A\cdot J_{(2\alpha/\beta)-2}\left\{2\sqrt{\frac{\alpha}{\mu}\left(\frac{2\alpha}{\beta}-1\right)}\right\} \tag{92}$$

When we note that the first term in this expression is just $\mu\ g_1(v)$ it is clear that (73) can be satisfied only when μ takes values such that

$$J_{(2\alpha/\beta)-2}\left\{2\sqrt{\frac{\alpha}{\mu}\left(\frac{2\alpha}{\beta}-1\right)}\right\} = 0. \tag{93}$$

Remembering, from (75), that $\mu = 4\lambda/\omega_0$ and writing[25] $\psi_0 = (\omega_0\alpha)/4$ for the total power output of the pre-multiplier filter, we see that (93) may be written

$$J_{(2\alpha/\beta)-2}\left\{2\sqrt{\frac{\psi_0}{\lambda}\left(\frac{2\alpha}{\beta}-1\right)}\right\} = 0. \tag{94}$$

We have shown earlier in the paper [see (39)] that the mth cumulant for the probability distribution of the output $x_0(t)$ may be expressed in terms of sums of the mth powers of the eigenvalues λ_j. Fortunately the computation of these sums is very readily carried out when we make use of an ingenious technique described by Spiegel[26] He has shown that, if x_j denotes the jth root of the equation (whose roots must all be real),

$$1 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3 + \cdots = 0, \tag{95}$$

then

$$\sum_{j=1}^{\infty}\frac{1}{x_j} = -\alpha_1 \tag{96}$$

$$\sum_{j=1}^{\infty}\frac{1}{x_j^2} = \alpha_1^2 - 2\alpha_2 \tag{97}$$

$$\sum_{j=1}^{\infty}\frac{1}{x_j^3} = 3\alpha_1\alpha_2 - 3\alpha_3 - \alpha_1^3 \tag{98}$$

$$\sum_{j=1}^{\infty}\frac{1}{x_j^4} = \alpha_1^4 - 4\alpha_1^2\alpha_2 + 2\alpha_2^2 + 4\alpha_1\alpha_3 - 4\alpha_4. \tag{99}$$

Now a theorem[27] due to Lommel states that all the zeros of $J_\nu(z)$ are real provided $\nu > -1$. Thus it follows that for $2\alpha/\beta - 1 > 0$, the only roots of (94) are real and hence Spiegel's method may be applied. Thus, expanding (94), we have

$$1 - \frac{\psi_0}{\lambda} + \frac{\left(\frac{2\alpha}{\beta}-1\right)}{2\left(\frac{2\alpha}{\beta}\right)}\frac{\psi_0^2}{\lambda^2} - \frac{\left(\frac{2\alpha}{\beta}-1\right)^2}{3!\left(\frac{2\alpha}{\beta}\right)\left(\frac{2\alpha}{\beta}+1\right)}\frac{\psi_0^3}{\lambda^3}$$
$$+ \frac{\left(\frac{2\alpha}{\beta}-1\right)^3}{4!\left(\frac{2\alpha}{\beta}\right)\left(\frac{2\alpha}{\beta}+1\right)\left(\frac{2\alpha}{\beta}+2\right)}\cdot\frac{\psi_0^4}{\lambda^4} - \cdots = 0, \tag{100}$$

and using (39) and (96) through (99) we find the following simple expressions for the first four cumulants:

$$K_1 = \sum\lambda_j = \psi_0 \tag{101}$$

$$K_2 = 2\sum\lambda_j^2 = \psi_0^2\cdot\frac{\beta}{\alpha} \tag{102}$$

$$K_3 = 2^3\sum\lambda_j^3 = \psi_0^3\cdot 8\cdot\frac{\beta}{\alpha}\cdot\frac{\beta}{2\alpha+\beta} \tag{103}$$

$$K_4 = 3\cdot 2^4\sum\lambda_j^4$$
$$= \psi_0^4\cdot 6\cdot\frac{\beta}{\alpha}\cdot\frac{\beta}{\alpha+\beta}\cdot\frac{\beta}{2\alpha+\beta}\cdot\frac{10\alpha+\beta}{\alpha}. \tag{104}$$

If we introduce the normalized variable

$$V = \frac{x_0(t) - K_1}{(K_2)^{1/2}} \tag{105}$$

then a convenient expansion for the probability distribution of V is given by the Edgeworth series[28]

$$p(V) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}V^2}\{1 + C_3H_3(V)$$
$$+ [C_4H_4(V) + C_6H_6(V)] + \cdots\}, \tag{106}$$

where $H_n(x)$ is the Hermite polynomial of degree n defined by

$$\frac{d^n}{dx^n}\{e^{-\frac{1}{2}x^2}\} = (-1)^n H_n(x)e^{-\frac{1}{2}x^2}, \tag{107}$$

[25] To show this, set $\tau = 0$ in (72).

[26] M. R. Spiegel, "The summation of series involving roots of transcendental equations and related applications," *Jour. Appl. Phys.*, vol. 24, p. 1103; September, 1953.

[27] G. N. Watson, *loc. cit.*, p. 482.

[28] H. Cramér, *loc. cit.*, p. 288.

nd where

$$C_3 = -\frac{1}{6}\frac{K_3}{(K_2)^{3/2}} = -\frac{4}{3}\frac{\left(\frac{\alpha}{\beta}\right)^{1/2}}{\frac{2\alpha}{\beta}+1} \tag{108}$$

$$C_4 = \frac{1}{24}\frac{K_4}{K_2^2} = \frac{5}{2}\frac{\left(\frac{\alpha}{\beta}+\frac{1}{10}\right)}{\left(\frac{2\alpha}{\beta}+1\right)\left(\frac{\alpha}{\beta}+1\right)}. \tag{109}$$

An explicit expression for c_6, which term in the Edge-vorth series is of the same order as the c_4 term, has not een given because expressions corresponding to (99) ut for powers greater than four are not yet available.[29]

It is clear from (106), (108) and (109) that for large α/β (the ratio of postmultiplier to premultiplier filter ime constants), the distribution of V tends to Gaussian, s we might expect from the central limit theorem. We ote that other expansions, corresponding to the Edge-vorth series, but more useful for intermediate values of α/β have been given by Marcum.[30]

To conclude our discussion of this example we note hat for the specific case $\alpha/\beta = 5/4$, (94) may be written

$$J_{1/2}\left\{2\sqrt{\frac{\psi_0}{\lambda}\cdot\frac{3}{2}}\right\} = 0. \tag{110}$$

When we recall that

$$J_{1/2}(x) = \sqrt{\frac{2}{\pi x}}\sin x, \tag{111}$$

t is clear that the eigenvalues are given by

$$\lambda_j = \frac{6}{j^2\pi^2}\cdot\psi_0, \qquad j = \cdots -2, -1, 0, 1, 2 \cdots. \tag{112}$$

We now use (112) in (35) and obtain, for the character-stic function

$$M(i\theta) = \prod_{j=1}^{\infty}\left\{1 - \frac{i\theta}{j^2\pi^2}\cdot(12\psi_0)\right\}^{-1}. \tag{113}$$

Thus the probability distribution $p(x_0)$ for the output $x_0(t)$ is given by

$$p(x_0) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\frac{e^{-i\theta x_0}}{\prod_{j=1}^{\infty}\left\{1 - \frac{i\theta}{j^2\pi^2}(12\psi_0)\right\}}\cdot d\theta. \tag{114}$$

The integrand has simple poles on the negative imagin-ry axis only and so evaluating the integral by contour integration we obtain[31]

$$p(x_0) = \frac{\pi^2}{6\psi_0}\sum_{s=1}^{\infty}(-1)^{s+1}s^2\exp -\left\{\frac{\pi^2 s^2}{12}\cdot\frac{x_0}{\psi_0}\right\}, \quad x_0 > 0. \tag{115}$$
$$= 0, \qquad x_0 < 0.$$

We now introduce the theta function $\theta_4(z; q)$ defined[32] by

$$\theta_4(z; q) = 1 + 2\sum_{n=1}^{\infty}(-1)^n q^{n^2}\cos 2nz \tag{116}$$

In particular, setting $z = 0$ we have:

$$\theta_4(0; q) = 1 + 2\sum_{n=1}^{\infty}(-1)^n q^{n^2} \tag{117}$$

and denoting differentiation with respect to q by a prime we find

$$q\theta_4'(0; q) = 2\sum_{n=1}^{\infty}(-1)^n n^2 q^{n^2} \tag{118}$$

Thus finally we may write (115) in the form

$$\left.\begin{aligned} p(x_0) &= \frac{\pi^2}{12\psi_0}\cdot\exp\left\{-\frac{\pi^2}{12\psi_0}\right\}\cdot x_0 \\ &\quad\cdot\theta_4'\left\{0; \exp\left\{-\frac{\pi^2}{12\psi_0}\right\}\cdot x_0\right\}, \qquad x_0 > 0 \\ &= 0, \qquad x_0 < 0 \end{aligned}\right\} \tag{119}$$

which is an explicit expression for the probability distributor in this case.

V. Conclusion

We have shown how the statistical problem considered in this paper may be solved in terms of a corresponding eigenvalue problem. Explicit solutions of the integral equations encountered have been given for a simple example.

We point out finally that a considerable amount of time has been spent trying to obtain explicit solutions of the integral equations for an example system involving a pure delay, as in Fig. 1, but so far little progress has been made. Of course the cumulants can be obtained in terms of the iterated kernels [derived by successive integration using (48)] as shown by Emerson[33] but the process is extremely tedious in practice.

VI. Acknowledgment

The author has benefited from the suggestions of many friends during the course of this work and in particular he wishes to acknowledge interesting discussions with Drs. A. J. F. Siegert, M. A. Meyer, and R. C. Emerson.

[29] In a private communication Dr. Spiegel advises that he hopes t some time to extend his results up to about the 10th power. lowever the computations are laborious.

[30] T. J. Marcum, "A statistical theory of target detection by ulsed radar," Mathematical Appendix, Rand Corp. Res. Memo. RM. 753, p. 35; 1952.

[31] For details of this evaluation see M. Kac and A. J. F. Siegert, *loc. cit.*

[32] E. T. Whittaker and G. N. Watson, "A Course of Modern Analysis," Cambridge University Press, p. 464, 1950.

[33] R. C. Emerson, *loc. cit.*

Interpolation and Extrapolation of Sampled Data

A. B. LEES†

Summary—This paper is essentially an extension of the optimum filtering theory of Zadeh and Ragazzini to the case where the time function to be operated upon is available only at a sequence of sample instants. As in the latter paper the signal is taken to consist of two parts, one being a polynomial $p(t)$ with unknown coefficients and known maximum degree n, and the other being a stationary random component $M(t)$ with known autocorrelation function. It is assumed that the signal has been contaminated before filtering by the addition of stationary random noise, also of known autocorrelation function, and that the input to the filter consists of a sequence of impulse functions of constant repetition period T, each impulse being of area equal to the sample value of the signal plus noise at the time of occurrence of the impulse. This paper shows how to find the weighting function, $W(t)$, of a linear filter which will convert the sequence of impulse functions into a smoothed output subject to the following conditions: the weighting function $W(t)$ is only nonzero over a finite range; in the absence of random components, the interpolation or extrapolation is error-free; in the presence of any random signal the approximation at the output follows a least squares law. The weighting function of the optimum filter is shown to be piecewise continuous in the intervals $\{rT \leq t \leq (r+1)T; r = 0, 1, 2, \cdots N\}$ and the paper concludes with a discussion of a simple example illustrating the practical application of the solution.

Introduction

THE FUNDAMENTAL problem of extracting, in a least squares sense, a signal $s(t)$ from a mixture $y(t)$ of signal and noise $n(t)$,

$$y(t) = s(t) + n(t) \tag{1}$$

by means of an optimum linear filter was first formulated and solved by Wiener [1] who considered the case where the signal and noise are members of stationary random processes with known auto- and cross-correlation functions and where the function $y(t)$ is available to the filter in the form of a continuous waveform. A subsequent paper of Zadeh and Ragazzini [2] extended the above theory to the case where the signal $s(t)$ contains, in addition to the random component $M(t)$, a polynomial $p(t)$:

$$p(t) = \sum_{r=0}^{n} a_r \left(\frac{t}{T}\right)^r \tag{2}$$

of known maximum order n, but with unknown coefficients $\{a_0, a_1, \cdots a_n.\}$, and where the linear filter employed is only required to have a finite memory.

In the practical development of sampled data systems, there arise some analogous problems [5–7], with the essential difference that information available about the input waveform is restricted to its values at discrete instants in time. The input waveform to the filter thus contains information relating to a periodic or quasi-periodic observation procedure on the original continuous waveform $s(t)$, and in practice this waveform approximates to a sequence of impulse functions, each impulse having an area equal to the ordinate of the original waveform $y(t)$ at the instant when the impulse occurs. If the sampling has been performed with a constant repetition period of duration T, then evidently such a sequence may be written

$$y_s(t) = \sum_{r=-\infty}^{\infty} y(-rT)\,\delta(t + rT), \tag{3}$$

and the problem discussed in this paper is the specification of the weighting function $W(t)$ of a linear filter having the following three fundamental properties:

1) In the absence of both random components, the filter output is to be just $p(t + \alpha)$ so that the input polynomial is reproduced with a time shift, but otherwise without error. If $\alpha > 0$ the operation is one of extrapolation; if $\alpha = 0$ the operation is instantaneous interpolation; if $\alpha < 0$ it is delayed interpolation.

2) In the presence of one or both of the random components the mean squared error over-all time between $s(t + \alpha)$ and the smoothed output of the filter is to be a minimum for the weighting function $W(t)$ satisfying condition 1.

3) The filter is required to have only a finite memory, implying that the weighting function $W(t)$ shall vanish for all time instants greater than some chosen number. In the interests of simplicity this number will be taken to be a multiple of the sampling interval T and the condition is written

$$W(t) = 0 \quad \text{for} \quad t > (N + 1)T,$$

where N is an integer.

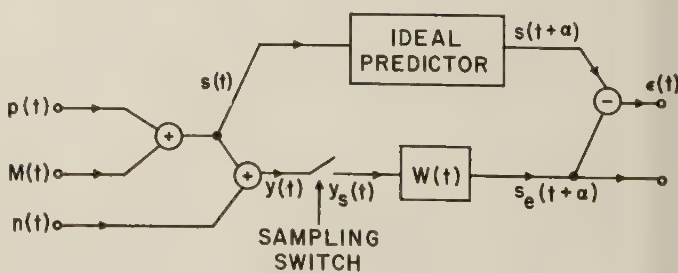


Fig. 1—System flow diagram.

A schematic representation of the various waveforms and of the system components pertinent to the following theory is given in Fig. 1.

† Visiting Professor of Electrical Engineering, Columbia University, New York, N. Y.

Since methods have already been described in the literature for approximating as closely as is desired to weighting functions of the type which arise in the succeeding analysis [3, 4], the specification of the optimum linear filter reduces to the specification of an optimum $W(t)$. The analytic procedure adopted may be divided briefly into three steps: to find the constraint equations imposed by the first and third conditions mentioned above; to obtain an expression for the mean squared error; and finally to find a $W(t)$ minimizing this latter subject to the constraint equations. The paper concludes with a discussion of a simple example illustrating the procedure.

Constraint Equations

Since the formulation of the problem has included no reference to a specific origin in time, it may be supposed, without any loss in generality, that the input to the filter in the absence of the random components may be written in the form

$$\sum_{r=-\infty}^{0} p(rT)\,\delta(t - rT) \tag{4}$$

and that the filter output is considered at some instant t in the interval $(0 \leq t < T)$. It has been stipulated at the outset that the polynomial $p(t)$ is of maximum degree n, which means that $p(t)$ can be represented everywhere in terms of its values at not more than $n + 1$ distinct points. Arguing from this conclusion, it is evident that when the filter is providing an output at time t, it must be operating on at least $n + 1$ previous input sample values for the output to be certainly free of error. This implies, in general, that the weighting function $W(t)$ must assume some values at least over the range $[0 \leq t < (n + 1)T]$, and in fact we shall assume that $W(t)$ is only required to vanish for $t > (N + 1)T$ where $N > n$. Since the response of a filter with weighting function $W(t)$ to an impulse $\delta(t)$ is just $W(t)$, the response to the input specified in (4), may be written

$$\sum_{r=0}^{-\infty} p(rT)W(t - rT),$$

and if the sign of the running index r is changed and this expression is equated to the required output $p(t + \alpha)$, the implied constraint may be expressed as

$$\sum_{r=0}^{\infty} p(-rT)W(t + rT) = p(t + \alpha) \tag{5}$$

for all $p(t)$ of degree n. The previous equation can be given a more convenient formulation by the introduction of a set of auxiliary functions $u_0(t)$, $u_1(t)$, $u_2(t) \cdots u_N(t)$, defined over the interval $0 \leq t < T$ and which are related to $W(t)$ in the following manner:

$$u_r(t) = W(t - rT); \;\; 0 \leq t < T; r = 0, 1, 2, \cdots N. \tag{6}$$

If these functions are used in (4) and the polynomials written in their expanded form by reference to (2), then an equation

$$\sum_{r=0}^{N} \sum_{s=0}^{n} a_s(-r)^s u_r(t) = \sum_{s=0}^{n} a_s\left[\frac{(t + \alpha)}{T}\right]^s$$

results, which must be true for all choices of $\{a_0, a_1, a_2 \cdots a_n\}$. This last statement implies that

$$\sum_{r=0}^{N} (-r)^s u_r(t) = \left(\frac{t + \alpha}{T}\right)^s \qquad (s = 0, 1, 2, \cdots n), \tag{7}$$

and the problem of finding constraints such that, in the absence of noise, the filter reproduces exactly the input polynomial, has been reduced to the selection of a set of functions $\{u_0(t),\ u_1(t),\ u_2(t)\ \cdots\ u_N(t)\}$ satisfying the $n + 1$ equations in (7). It may be noted here that by inspection of (7), no general solution exists if $N < n$, confirming the argument given earlier in this section, that for error-free reproduction of the input polynomial, the filter must be able to remember at least $n + 1$ previous sample values. If N is equal to n, there is, in general, just one set of u's satisfying (7), and these define the weighting function with the shortest memory which will accurately reproduce the polynomial; the fact that there is only one $W(t)$ in this case naturally precludes any optimization procedure in the presence of noise. Finally, it will be observed that if $N > n$, there is, in general, an infinity of solutions to (7), and the optimization problem will be concerned with finding some set of u's out of this infinity of possibilities corresponding to a minimum mean squared error in the presence of noise.

Mean Squared Error

In the case where the composite signal consisting of a polynomial $p(t)$ and a stationary random waveform $M(t)$ has been modified by the addition of stationary random noise $n(t)$, the input to the linear filter may be written as

$$y_s(t) = \sum_{r=-\infty}^{\infty} [p(rT) + M(rT) + n(rT)]\delta(t - rT),$$

and the estimate of $s(t + \alpha)$ at the output of the filter becomes

$$s_e(t + \alpha) = \sum_{r=0}^{\infty} [p(-rT) + M(-rT) + n(-rT)]W(t + rT),$$

by a modification of (5). Introducing now the functions $u_r(t)$ and assuming that they have been constrained to satisfy (7) the error $\epsilon(t)$ between $s_e(t + \alpha)$ and $s(t + \alpha)$ may be written

$$\begin{aligned}\epsilon(t) &= s(t + \alpha) - s_e(t + \alpha)\\ &= M(t + \alpha) - \sum_{r=0}^{N} [M(-rT) + n(-rT)]u_r(t).\end{aligned} \tag{8}$$

It may be recalled that the problem was formulated for interpolation instants such that $0 \leq t < T$ and evaluating

the mean squared error over this interval yields,

$$\langle\epsilon^2(t)\rangle_T = \frac{1}{T}\int_0^T M^2(t+\alpha)\,dt$$
$$+\frac{1}{T}\sum_r^N\sum_s^N [M(-rT)+n(-rT)]$$
$$\cdot [M(-sT)+n(-sT)]\int_0^T u_r u_s\,dt$$
$$-\frac{2}{T}\sum_r^N [M(-rT)+n(-rT)]\int_0^T M(t+\alpha)u_r(t)\,dt \quad (9)$$

If the above averaging procedure is repeated over all intervals of length T such that $[rT \le t < (r-1)T]$, where r ranges from $-\infty$ to $+\infty$ and an average taken of all these results, then this is evidently just the same as performing one averaging operation over the range $(-\infty < t < \infty)$. With a little algebra it may be shown that this result is

$$\langle\epsilon^2(t)\rangle_\infty = \langle M^2(t)\rangle + \frac{1}{T}\sum_r^N\sum_s^N \{\phi_M[(r-s)T]$$
$$+\phi_n[(r-s)T]\}\int_0^T u_r u_s\,dt$$
$$-\frac{2}{T}\sum_r^N\int_0^T \phi_M[t+\alpha+rT]u_r(t)\,dt, \quad (10)$$

where $\phi_M(\tau)$ is the autocorrelation function of the signal process $M(t)$; $\phi_n(\tau)$ is the autocorrelation function of the noise process $n(t)$; and $\langle M^2(t)\rangle$ is the mean power of the waveform $M(t)$. In deriving the above expression it has been assumed that both of the random functions $M(t)$ and $n(t)$ have zero means and that they are linearly independent. Otherwise additional terms appear in the resulting equation for the mean squared error.

Minimization Procedure

On the basis of the preceding formulation the problem remaining is to choose the set of functions $\{u_r\}$ to be that particular set satisfying the constraints of (7) which minimizes the right hand side of (10). For convenience and brevity it will be assumed in the following analysis that the sampling repetition interval T is unity. This involves no loss of generality since the optimum $W(t)$ obtained on this assumption may be generalized by replacing t by t/T and on the basis of this assumption the equations relevant to the minimization problem may now be summarized as,

$$\langle\epsilon^2(t)\rangle_\infty = \langle M^2(t)\rangle$$
$$+\sum_r^N\sum_s^N \{\phi_M(r-s)+\phi_n(r-s)\}\int_0^1 u_r u_s\,dt$$
$$-2\sum_r^N\int_0^1 \phi_M(t+\alpha+r)u_r(t)\,dt \quad (11)$$

and

$$\sum_{r=0}^N (-r)^s u_r(t) = (t+\alpha)^s \qquad (s = 0, 1, 2, \cdots n). \quad (12)$$

The constraint equations have to be true for all t in $(0 \le t < 1)$ and since the right-hand sides of these equations are functions of time, then on introducing the undetermined multipliers $\{\lambda_0, \lambda_1, \cdots, \lambda_n\}$ which will be themselves functions of time, and upon noting that $\langle M^2(t)\rangle$ is a constant, the problem is transformed into the minimization of the expression,

$$\sum_s^N\sum_r^N \{\phi_M(r-s)+\phi_n(r-s)\}\int_0^1 u_r u_s\,dt$$
$$-2\sum_r^N\int_0^1 \phi_M(t+\alpha+r)u_r\,dt$$
$$-\int_0^1\sum_{k=0}^n \lambda_k(t)\sum_r^N (-r)^k u_r(t)\,dt. \quad (13)$$

To obtain a more convenient notation define the functions ψ_{rs} and $\theta_r(t)$ as

$$\psi_{rs} = \phi_M(r-s)+\phi_n(r-s) \quad (14)$$

$$\theta_r(t) = 2\phi_M(t+\alpha+r), \quad (15)$$

and now perturbing the $\{u_r\}$ by the addition of arbitrary functions $\xi_r(t)$ in the interval $(0 \le t < 1)$, the variation of (13) becomes

$$\sum_s^N\int_0^1 \xi_r(t)\left\{\sum_r^N \psi_{rs}u_r(t) - \sum_{k=0}^n \lambda_k(t)(-s)^k - \theta_s(t)\right\}dt$$

and this will be always zero for all choices of $\{\xi_r\}$, provided that the $\{u_r(t)\}$ satisfy the $N+1$ simultaneous equations

$$\sum_{r=0}^N \psi_{rs}u_r(t) = \theta_s(t) + \sum_{k=0}^n (-s)^k\lambda_k(t). \quad (16)$$

If ψ_{rs}^{-1} is a typical element of the inverse of the matrix $[\psi]$, then the solution for u_r may be written explicitly as

$$u_r(t) = \sum_{s=0}^N \psi_{rs}^{-1}\left\{\sum_{k=0}^n (-s)^k\lambda_k(t) + \theta_s(t)\right\},$$
$$(r = 0, 1, 2, \cdots N). \quad (17)$$

and all that remains is to evaluate the undetermined multipliers from the $n+1$ constraint equations (12); thus the $\lambda_k(t)$ satisfy

$$\sum_{r=0}^N (-r)^s\sum_{m=0}^N \psi_{rm}^{-1}\left\{\sum_{k=0}^n (-m)^k\lambda_k(t) + \theta_m(t)\right\} = (t+\alpha)^s$$
$$(s = 0, 1, 2, \cdots n)$$

or by rearrangement

$$\sum_{k=0}^n \lambda_k(t)U_{ks}^{\,1} = (t+\alpha)^s - \sum_{r=0}^N\sum_{m=0}^N (-r)^s\psi_{rm}^{-1}\theta_m(t), \quad (18)$$

where the element of the inverse matrix of $[U]$ is

$$U_{ks}^{-1} = \sum_{r=0}^{N} \sum_{m=0}^{N} (-r)^s (-m)^k \psi_{rm}^{-1}. \qquad (19)$$

Thus using (18) the multipliers λ_k have the explicit representation

$$\lambda_k(t) = \sum_{s=0}^{n} U_{sk}(t + \alpha)^s - \sum_{r=0}^{N} \sum_{m=0}^{N} \sum_{s=0}^{n} (-r)^s \psi_{rm}^{-1} U_{sk} \theta_m(t), \qquad (20)$$

and this together with (17) provides the complete analytical solution to the problem of specifying the optimum set $\{u_r\}$ and thus to the problem of specifying the weighting function $W(t)$ of the optimum linear filter. Provided that the functions $\theta_r(t)$ are continuous in the interval $(0 \leq t < 1)$, which from a physical point of view seems reasonable, then the members of the set $\{u_r\}$ will be continuous and thus $W(t)$ itself is piecewise continuous in the intervals $\{0, 1\}, \{1, 2\} \cdots \{N, N + 1\}$.

In certain special cases of practical importance the solution obtained above assumes a simpler form and consideration will now be given to two such cases. The input signal $s(t)$ was assumed in the beginning to consist of two parts $M(t)$ and $p(t)$ and let attention be restricted now to the case where $M(t)$ is identically equal to zero. Then following through the arguments of the preceding sections it is evident that the constraint equations, which do not depend on the presence or absence of an $M(t)$, will be unchanged, and thus the mean squared error over the infinite interval $(-\infty < t < \infty)$ has the form

$$\langle \epsilon^2(t) \rangle_\infty = \sum_{r}^{N} \sum_{r}^{N} \phi_n(r - s) \int_0^1 u_r(t) u_s(t) \, dt. \qquad (21)$$

From this we deduce immediately that the θ_r's can be made zero in the final solutions in (17) and (20) yielding

$$u_r(t) = \sum_{s=0}^{N} \psi_{rs}^{-1} \sum_{k=0}^{n} (-s)^k \lambda_k(t) \qquad (r = 0, 1, \cdots N), \qquad (22)$$

$$\lambda_k(t) = \sum_{s=0}^{n} U_{sk}(t + \alpha)^s \qquad (k = 0, 1, \cdots n), \qquad (23)$$

where in this case

$$\psi_{rs} \triangleq \phi_n(r - s). \qquad (24)$$

Combining (22) and (23) gives an expression for $u_r(t)$,

$$u_r(t) = \sum_{s=0}^{N} \psi_{rs}^{-1} \sum_{k=0}^{n} (-s)^k \sum_{m=0}^{n} U_{mk}(t + \alpha)^m = \sum_{m=0}^{n} K_{rm}(t + \alpha)^m, \qquad (25)$$

where the matrix K_{rm} is defined by

$$K_{rm} = \sum_{s=0}^{N} \sum_{k=0}^{n} (-s)^k U_{mk} \psi_{rs}^{-1}. \qquad (26)$$

It may be observed that if (26) is multiplied by $(-r)^p$ and summed over r from 0 to N, then on using (19), an equation

$$\sum_{r=0}^{N} K_{rm}(-r)^p = \sum_{k=0}^{n} U_{mk} U_{kp}^{-1} \qquad (27)$$

results. Now if the summation over k on the right-hand side had been from 0 to N instead of from 0 to n, then this would have yielded the Kronecker symbol δ_{mp} indicating that $[K]$ has to be the matrix inverse to $[(-r)^p]$. As would be anticipated this is just the solution for the $u_r(t)$ in the case where $n = N$ and only one set $\{u_r\}$ exists satisfying the constraint equations. In the general case the right-hand side of (27) defines a nondiagonal matrix depending on the correlation matrix $[\psi]$ and the solution is no longer so simple to obtain. It may be noted finally that in this special case where $M(t) \leq 0$, the solution in (25) defines an optimum weighting function $W(t)$ which is piecewise continuous in the intervals $(0, 1), (1, 2) \cdots (N, N + 1)$ in the form of polynomials of degree n.

The second special case of the general solution, (17) and (20), which will be considered, is where the noise amplitudes at successive sampling instants are linearly independent, and where it is assumed again that $M(t) \triangleq 0$. In this case the correlation matrix $[\phi_n(r - s)]$ of the noise amplitudes is diagonal; *i.e.*,

$$\phi_n(r - s) = \begin{cases} \sigma^2 & (r = s) \\ 0 & (r \neq s), \end{cases} \qquad (28)$$

where σ^2 is the noise power. If this result is used in (22) then the expression for $u_r(t)$ may be written

$$u_r(t) = \frac{1}{\sigma^2} \sum_{k=0}^{n} (-r)^k \lambda_k(t) \qquad (r = 0, 1, \cdots N), \qquad (29)$$

where now, the $\lambda_k(t)$ satisfy equations

$$\sum_{k=0}^{n} \lambda_k(t) \sum_{r=0}^{N} (-r)^{s+k} = \sigma^2(t + \alpha)^s \ (s = 0, 1, \cdots N). \ (30)$$

Defining the matrix Ω_{sk} by

$$\Omega_{sk}^{-1} = \sum_{r=0}^{N} (-r)^{s+k} \qquad \left(\begin{matrix} s \\ k \end{matrix} = 0, 1, \cdots N\right), \qquad (31)$$

the solution for $\lambda_k(t)$ becomes, from (30)

$$\lambda_k(t) = \sigma^2 \sum_{s=0}^{n} \Omega_{sk}(t + \alpha)^s. \qquad (32)$$

This last summation may be computed, using (31), to find the numerical values of the coefficients in the polynomials $\{u_r(t)\}$.

Example

As an example of the use of the analytical result derived above, consider the case of a linear predictor with the following special conditions; $M(t) = 0$; $n = 1$; and with the noise uncorrelated between sampling instants.

Then from (31) the matrix $[\Omega]^{-1}$ has the form

$$[\Omega]^{-1} = \begin{bmatrix} (N+1) & -\frac{N(N+1)}{2} \\ -\frac{N(N+1)}{2} & \frac{N(N+1)(2N+1)}{6} \end{bmatrix}, \tag{33}$$

and upon inversion of this, the matrix $[\Omega]$ may be written

$$[\Omega] = \begin{bmatrix} \frac{2(2N+1)}{(N+1)(N+2)} & \frac{6}{(N+1)(N+2)} \\ \frac{6}{(N+1)(N+2)} & \frac{12}{N(N+1)(N+2)} \end{bmatrix}. \tag{34}$$

Using this result in (32) yields

$$u_r(t) = \frac{2(2N+1)}{(N+1)(N+2)} - \frac{6r}{(N+1)(N+2)} + (t+\alpha)\left\{\frac{6}{(N+1)(N+2)} - \frac{12r}{N(N+1)(N+2)}\right\}. \tag{35}$$

For the special case of $N = 2$;

$$u_r(t) = \frac{5}{6} - \frac{r}{2} + (t+\alpha)\left\{\frac{1}{2} - \frac{r}{2}\right\} \tag{36}$$

and the functions become explicitly

$$\begin{aligned} u_0(t) &= \frac{5}{6} + \frac{1}{2}(t+\alpha) \\ u_1(t) &= \frac{1}{3} \\ u_2(t) &= -\frac{1}{6} - \frac{1}{2}(t+\alpha). \end{aligned} \tag{37}$$

The corresponding weighting function is shown in Fig. 2(b) and may be compared with the form of the "shortest—memory" weighting function Fig. 2(a). For the special case $N = 3$, a similar procedure yields the set $\{u_r(t)\}$

$$u_r(t) = \frac{7}{10} - \frac{3r}{10} + (t+\alpha)\left\{\frac{3}{10} - \frac{r}{5}\right\}, \tag{38}$$

or more explicitly, [Fig. 2(c)]

$$\begin{aligned} u_0(t) &= \frac{7}{10} + \frac{3}{10}(t+\alpha) \\ u_1(t) &= \frac{4}{10} + \frac{1}{10}(t+\alpha) \\ u_2(t) &= \frac{1}{10} - \frac{1}{10}(t+\alpha) \\ u_3(t) &= -\frac{2}{10} - \frac{3}{10}(t+\alpha). \end{aligned} \tag{39}$$

It is of interest to compare the mean squared error in each of these two cases with that for the case of $N = 1$. In this latter case the $u_r(t)$ are

$$\begin{aligned} u_0(t) &= 1 + (t+\alpha) \\ u_1(t) &= -(t+\alpha), \end{aligned} \tag{40}$$

using (12); and substituting these expressions in the equation for the mean squared error

$$\langle \epsilon^2(t) \rangle_\infty = \sigma^2 \sum_r^N \int_0^1 u_r^2(t)\,dt,$$

which is a modification of (21), using condition (28), the result becomes

$$\langle \epsilon^2(t) \rangle_\infty = \sigma^2\left[\frac{8}{3} + 2\alpha^2 + 4\alpha\right]. \tag{41}$$

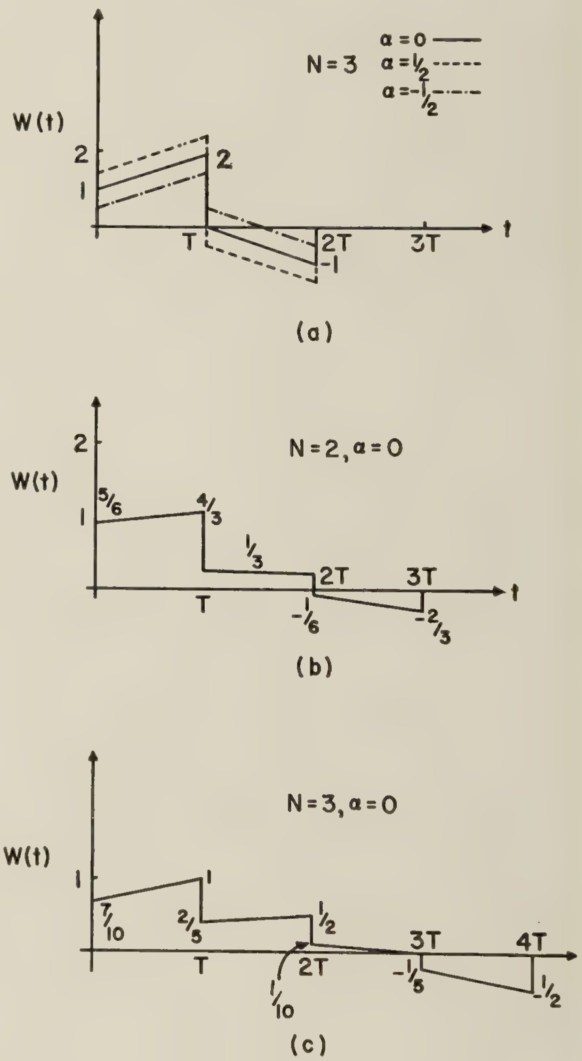


Fig. 2—Weighting functions for straight-line prediction with $N = 1$, 2, and 3.

In the case $N = 2$, an analogous calculation shows that

$$\langle \epsilon^2(t) \rangle_\infty = \sigma^2\left[\frac{3}{2} + \frac{\alpha^2}{2} + \frac{3\alpha}{2}\right], \tag{42}$$

and for $N = 3$,

$$\langle \epsilon^2(t) \rangle_\infty = \sigma^2\left[\frac{16}{15} + \frac{\alpha^2}{5} + \frac{4\alpha}{5}\right]. \tag{43}$$

Evidently the mean squared error is reduced for all values of α, by the successive increases in the value of N; in fact, for instantaneous interpolation, where $\alpha = 0$, the mean squared error is reduced by factors 0.5625 and 0.4 from its value at $N = 1$, for values of N equal to 2 and 3 respectively.

Conclusion

It will have been observed from the nature of the analytical solution to the optimization problem, that the weighting function $W(t)$ is finite and piecewise continuous over its range of definition. Unlike the optimum weighting functions of Zadeh and Ragazzini, it does not involve any impulse functions. The presence of such singularities would imply a direct transmission through the filter (with a time shift) of the input waveform and since this latter consists of impulse functions, then the filter output itself would contain impulse functions. It was taken to be axiomatic that in a smoothed output, singularity functions would be undesirable. On the other hand the optimum weighting function contains, in general, discontinuous jumps at the points 0, T, $2T$, $\cdots$, $(N + 1)T$ and this implies that the synthesis should be active. To find an optimum $W(t)$ which is continuous in value over the range $[0 \leq t < (N + 1)T]$ would require the introduction of $(N - 1)$ further constraints.

Bibliography

[1] Wiener, N., *The Interpolation, Extrapolation and Smoothing of Stationary Time Series.* New York, John Wiley & Sons, Inc., 1949.

[2] Zadeh, L. A., and Ragazzini, J. R., "An Extension of Wiener's Theory of Prediction." *Journal of Applied Physics*, Vol. 21 (July, 1950), pp. 645–655.

[3] Porter, A., and Stoneman, F. W., "New Approach to the Design of Pulse Monitored Servo Systems." *Proceedings of the Institution of Electrical Engineers*, part II, Vol. 97 (1950), p. 597.

[4] Bergen, A. R., and Ragazzini, J. R., "Sampled-Data Processing Techniques for Feedback Control Systems," *Transactions of the American Institute of Electrical Engineers*, "Applications and Industry." (November, 1954), pp. 1–12.

[5] Franklin, G., *The Optimum Synthesis of Sampled-Data Systems*, Technical Report T-6/B, Columbia University Electronic Research Laboratories.

[6] Barker, R. H., *Theory of Pulse Monitored Servomechanisms and their Use for Prediction*, Report No. 1046, Signals Research and Development Establishment. (England).

[7] Brogan, J., "Filters for Sampled Signals," *Proceedings of the Information Theory Symposium* (April, 1954), Brooklyn Polytechnic Institute.

The Third London Symposium on Information Theory*

NELSON M. BLACHMAN†

AT THE FIRST London Symposium on Information Theory, held at the Royal Society in September, 1950, about twenty papers were presented, including six on applications of information theory to psychology and neurophysiology. At the second symposium, held at the Institution of Electrical Engineers in September, 1952, emphasis on the latter fields was replaced by an emphasis on the transmission and analysis of speech, with which eight of the thirty-eight papers dealt. The scope of the third London Symposium on Information, held at the Royal Institution in September, 1955, was broadened to include not only all of these fields but the mechanical translation of languages as well, in addition to the basic topics of information theory, coding, etc. The breadth of its scope is also indicated by the range of backgrounds of the participants, which included: anatomy, animal welfare, anthropology, computers, economics, electronics, linguistics, mathematics, neuropsychiatry, neurophysiology, philosophy, phonetics, physics, political theory, psychology, and statistics. Of the 250 participants, half were British; one-eighth came from the United States (Shannon was not among them); the remainder, in order of decreasing numbers, were from the Netherlands, Sweden, France, Germany, Denmark, the USSR, Italy, Belgium, Switzerland, Spain, and Israel.

The auditorium of the Royal Institution, at 21 Albemarle Street, between Berkeley Square and Picadilly Circus, suited the size and type of the meeting perfectly, being semicircular in shape, with steeply rising tiers of seats, so that everyone in the room could see and hear everyone else. The program of the symposium was arranged to cover five days, Monday to Friday, September 12 to 16, with about seven papers presented each day; three in the morning, with a 15-minute coffee interval between the second and third; and four in the afternoon, with a half-hour tea interval in the middle. In this way, there was always a break just before or just after each paper, and possibility of tedium was easily avoided. Forty-five minutes were allotted to each paper, of which fifteen minutes were for the presentation and the remainder

* The author attended this symposium under the auspices of the U. S. Army Signal Corps.

† Electronic Defense Lab., Mountain View, Calif.

for discussion which generally proved quite lively. This arrangement, coupled with the prior distribution of mimeographed copies of all papers to the participants, proved quite felicitous, and Professor E. C. Cherry of the Imperial College of Science and Technology, who organized the symposium, deserves to be complimented for his efforts.

Monday: Information Theory

Professor B. van der Pol, Director of the CCIR (International Radio Consultative Committee), Geneva, very ably chaired the first day's sessions of the symposium. He started it off with a reference to his own studies of the effect of successive translation, in particular from English to French to English to French to English, each time by a different translator. He found that the meaning of the original text was retained through the translations but the style of the writing was lost, being replaced to a certain extent by the translators' style.

The first talk on the program was that of Professor J. L. van Soest of the Hague on "Some Consequences of the Finiteness of Information." He discussed ways of ensuring that the entropies of continuous distributions be non-negative.

He was followed by G. Spencer Brown of Christ Church College, Oxford University, who talked on "Chance and Control: Some Implications of Randomization." Brown called attention to some difficulties in the generation and testing of random sequences. He also pointed out that in the experimental testing of hypotheses, the statistical significance of the result is not a measure of the amount of information obtained. A heated discussion followed between Brown and I. J. Good of Cheltenham, Good claiming that a finite sequence must be regarded as random if the hypothesis is not rejected that it is part of an infinite random sequence. Good also pointed out that to every real number x between 0 and 1 corresponds an infinite sequence, *e.g.*, the decimal or binary expansion of x, and that for nearly all x this sequence is random.

The third paper, "On Some Measures of Information Used in Statistics," by M. P. Schuetzenberger of Paris, used lattice theory to discuss measures of information.

The afternoon session of the first day began with "Optical Transmission," by D. Gabor of the Imperial College of Science and Technology, London. Dr. Gabor discussed the transmission of light from one aperture (the input) to another aperture (the output). Here Shannon's $2WT$-dimensional vector describing a signal voltage is replaced by a Hermitian illumination matrix. The diagonalization of this matrix corresponds to the resolution of the optical field into its separate incoherent components. Each such component corresponds to one of Shannon's vectors. In the astronomical case, the focusing of a telescope at infinity effects this diagonalization, resolving the field into its individual incoherent point sources, the stars.

The next paper, "Two-dimensional Photographic Autocorrelation of Pictures and Alphabet Letters," by W. Meyer-Eppler and G. Darius of the Institut fuer Phonetik und Kommunikationsforschung der Universitaet Bonn, dealt with a very interesting and simple mechanism for the display of the cross-correlograms of pairs of pictures. One transparency, backed by ground glass, is used as a source of illumination, the light passes through a smaller transparency at some distance, and the correlogram is produced on a ground-glass screen at an appropriate distance beyond. The paper discussed the symmetry characteristics of the correlograms and their possible application to a reading machine.

J. F. Schouten of the Philips Telecommunication Industries, Hilversum, the Netherlands, then presented a paper on "Ignorance, Knowledge, and Information," which discussed the concept of "irrelevant information" and suggested that the sum of the measures of ignorance and information should be a constant during any communication process. In the discussion that followed, I. J. Good suggested that the random variable X be distinguished from x, any value which this variable may take on, in order to distinguish averaged conditional entropies from those which have not been averaged over X. He also suggested a notation for the amount of information about X which is contained in Y when Z is known to equal z, viz., $I(X{:}Y/z)$. Yehoshua Bar-Hillel of M.I.T. took exception to an example Schouten had given in which information had reduced knowledge and increased ignorance. He objected to this paradoxical use of familiar terms, and he pleaded, as he was to do several times during the symposium, for "semantic hygiene." Schouten's situation was one, he said, of "less certainty," rather than "less knowledge."

The last talk on the first day, a very interesting one, dealt with "Strategic Information in Games and in Voting," by Robin Farquharson of Nuffield College, Oxford. Farquharson discussed the difficulties which can result from the voter's inability to express all of his preferences in the casting of a single vote. Learning afterward how the others voted, one may consequently wish one had voted differently. The situation in which no one would want to change his vote after finding out how the others voted is called equilibrium. According to game theory, whenever voting is open, an equilibrium point exists such that no voter need regret his choice of voting strategy. However, perfect information about others' votes makes possible the bandwagon and underdog effects, which may yield equilibria without giving the voters their first choice that might also have been an equilibrium. A secret ballot is therefore necessary in order that voters may have strategies which are unconditionally best, for, if other voters are free to respond either way to one's own vote, then, whatever strategy one may choose, one may prove to have made the wrong choice.

Tuesday: Coding, Taxonomy, Etc.

Professor Arthur E. Laemmel of the Polytechnic Institute of Brooklyn started off the second day of the symposium with a talk on "A General Class of Codes and their Physical Realization," emphasizing the "general-purpose" realization of coding through the use of a multiple-track magnetic drum on which are recorded pulses that control the sequence of coding operations.

He was followed by Peter Elias of M.I.T., who talked on "Coding for Two Noisy Channels," namely, the "binary symmetric channel," in which each of the two possible transmitted symbols has the same probability of being received correctly and the complementary probability of being mistaken for the other, and the "binary erasure channel," which is similar except that errors are received as blanks (a third symbol) rather than as the incorrect symbol. Elias has found that when transmitting at rates much below the channel capacity, it pays to be clever about coding rather than to use Shannon's random code. In particular, he has found that check-symbol codes are as good as any other kind, in terms of both maximum transmission rate and error probability.

Elias has not studied particular codes, but David Slepian of the Bell Telephone Laboratories has studied certain codes for the symmetric binary channel in considerable detail. Although it was not on the program of the symposium, his very interesting paper "A Class of Binary Signaling Alphabets" will evidently be incorporated into the proceedings. Slepian starts with an "alphabet" of 2^k n-binary digit "letters" which form a group under addition modulo 2. With each of these letters he associates 2^{n-k} different n-binary-digit symbols which are at least as close to it (in the agreement of digits) as they are to any other letter. Each associated set is obtained as a coset of the corresponding letter. In this way, he is able to construct a maximum-likelihood detector. The probability of correct detection can be maximized by an appropriate choice of the alphabet, for the determination of which Slepian has compiled an extensive table. In this alphabet, there are k binary digits, fixed in position, giving information, and $n - k$ parity-check digits, which are the sums modulo 2 of fixed subsets of the information digits.

D. A. Huffman of M.I.T. was next on the program with an excellent paper on "The Synthesis of Linear Sequential Coding Networks" out of modulo-2 adders and delays of fixed size. Because of the possibility of having feedback paths within such networks, their outputs are, in effect, parity checks over an infinite past range of network inputs. Their transfer characteristics may be described by rational functions of the delay operator D. Huffman has studied the meaning of the "natural frequency" and the steady-state and transient responses of such networks and has developed methods for their realization using the least possible number of delay elements. When such a network incorporates feedback, it is capable of delivering an output with no input (*i.e.*, a sequence of zeros as input). Such an output, which is necessarily periodic, is called a "null sequence" of the inverse network, since the network having the reciprocal transfer characteristic would yield a sequence of zeros as output with this null sequence as input. When its transfer characteristic is a polynomial of degree n in D, the longest possible null-sequence length is $2^n - 1$. Such a maximum-length null sequence, interestingly, has properties somewhat like those of random noise.

Next came "An Elementary Proof of a Special Case of the Coding Theorem" by Professor G. A. Barnard of Imperial College, London, who proved the coding theorem for the symmetric binary channel. During the discussion period following Barnard's paper, S. K. Zaremba of Boulton Paul Aircraft Limited, Wolverhampton, introduced his own "Note on the Fundamental Theorem for a Discrete Channel with Noise," which asserts, on what appear to be dubious grounds, that Shannon's proof of this theorem is not correct.

Dr. Lars-Henning Zetterberg of Stockholm then talked on "A Comparative Study of Delta and Pulse-Code Modulation." "Delta modulation" involves the transmission of either a positive or a negative pulse (of the same size) at regular intervals, accordingly as the sum of all the pulses previously transmitted is less or greater than the present level of the modulating signal. Zetterberg compared single-channel delta and pulse-code modulation from the point of view of information theory and of the quality of transmission of speech-like signals in the absence of noise. He found that both systems have about the same channel capacity when their numbers of amplitude levels are the same and are large. When parameters are chosen to give a maximum ratio of signal to overloading and granulation noise, he found that delta modulation needs a much larger bandwidth than PCM if the desired ratio is moderate or high, but it becomes relatively more favorable when the bandwidth of the speech signals is large.

Next Mr. R. A. Fairthorne of the Royal Aircraft Establishment, South Farnborough, talked on "Some Clerical Operations on Languages," which dealt with marshaling as well. He was followed by Calvin N. Mooers of the Zator Company, Boston, who discussed "Information Retrieval upon Structured Content" for the cataloging of documents. He suggested the use of interlocking sets of descriptors, to indicate the interrelation of the descriptors, each of these sets possibly being designated by the superposition of corresponding random codes. The practicability of such methods, however, has not yet been ascertained.

Wednesday: Language Analysis and Mechanical Translation

The third day began with "Negative Entropy of Welsh Words," by Dr. D. A. Bell and Professor Alan S. C. Ross of the University of Birmingham. This paper ostensibly concerned the redundancy in the formation of English

words of n letters and of Welsh words of n phonemes. In both cases, the redundancies were determined by counting n-letter or n-phoneme words in a dictionary. In the case of English, the results were then weighted according to the frequency of n-letter words, and a mean redundancy per letter of English was thereby determined. Since a list of word frequencies for Welsh was unavailable, it was not possible to commit this same error for Welsh. During the discussion period, quite a number of objections were raised.

The next paper was "Mathematical Theory of Word Formation," by Professor W. Fucks of the Physikalisches Institut, Technische Hochschule, Aachen. He discussed the empirical distribution of the number of syllables per word in English, Arabic, German, Esperanto, Greek, Japanese, Russian, Turkish, and Latin, to which he was able to fit Poisson distributions shifted one unit upward so that no word has less than one syllable. During the discussion period the point was raised that Russian has several words of no syllables. Fucks had no theory to explain his success with Poisson distributions, but he exhibited a sort of pin-ball machine which is able to approximate this distribution.

Professor Benoit Mandelbrot of the Universities of Geneva and Grenoble then discussed "Two Statistical Laws of Language," which included "A Law of Berry and the Definition of Stress," dealing with the location of the syllabic and word stresses, and "A Thermodynamical Study of Systems of Categories with Willis (Natural) Structures." The Willis distribution, which states that the proportion of objects falling within the n-th largest category is $bn^{-(1+a)}$, with a between 0 and 1, according to Mandelbrot, can be considered as an approximation to an exceptional stable distribution of Cauchy-Paul Lévy. His study is said to parallel that of the normal stable distribution of Laplace-Gauss that arises in thermodynamics. During the discussion period, Professor A. S. C. Ross, a linguistician, indicated his own disagreement with the Zipfian law of word frequencies, which asserts the applicability of the Willis distribution. Marcel Golay of SCEL and Professor van der Pol suggested the information-theoretical study of music. It was then reported that random music in the styles of Bach and Vivaldi has already been generated at the Bell Telephone Laboratories. Henry Quastler of the University of Illinois claimed that Texas cowboy songs can be generated by a first-order Markoff process, *i.e.*, by a random process that depends upon only the previous note, supplemented by words chosen randomly according to certain frequencies from an appropriate list.

The next paper, by Professors S. Ceccato and E. Maretti of Milan, was titled "The Construction of a Translating Machine." This paper dealt with the difficulties confronting the designer of a translating machine, particularly if he wants to be sure that ideas get across correctly. Notation was presented which may have been intended to assist in the translation of ideas, but, perhaps because no mechanical apparatus is presently available to translate Italian into English, its significance was lost. Afterward, Bar-Hillel suggested it is hopeless to try to translate "ideas" rather than language symbols, since authors' ideas are often hard to pin down.

Two somewhat related papers were then presented together: "Influence of Context upon Translation," by Dr. A. D. Booth of Birkbeck College, London, and "A Program for Braille Translation," by J. P. Cleave, which dealt with the programming for an automatic digital computer—presumably the $APE(X)C$—of the transcription of English into Braille. This problem is of interest because Braille uses a number of contractions in ways that are a little complicated to describe; *i.e.*, the transcription is affected by the context. Dr. Booth demonstrated that only 14 comparisons are necessary to locate a given word in the Concise Oxford Dictionary; binary searching through a translating-machine's dictionary should thus be much faster than comparing with all entries in sequence. In the discussion period, Professor van der Pol pointed out a very interesting example of the influence of context upon translation—a nearby street sign reading ST. GEO. ST. Fairthorne suggested that the process of translation, translation back, correction of errors, retranslation, etc., until no errors remained would resemble the search for the eigenvalues of a matrix (or operator).

Following the afternoon tea interval, Professor Victor Yngve of M.I.T. talked on "The Translation of Languages by Machine." At M.I.T. progress is slowly being made on the mechanical translation of German into English; machine dictionaries still need to be set up. The machine, knowing the grammars of the two languages and thus being able to recognize types of words, such as parts of speech, is to code the input message into a transition (intermediate) language having little redundancy, then decode it into the target language subject to the constraints of the latter. This intermediate language might be useful in a secrecy system. Margaret Meade suggested, during the discussion period, that pidgin English, whose vocabulary is 75 per cent English and whose grammar is generalized Melanesian, might be a good language to work on, since it exhibits the important problems of mechanical translation.

The last talk of the day was a very lucid one on "Experiments in Mechanical Speech Recognition," by Drs. D. B. Fry and P. Denes of University College, London. Their recognizer uses about twenty $\frac{1}{3}$-octave filters in the range from 125 cps to 8 kc, computes the matrix of cross-products of pairs of filter outputs, and determines the most likely of 15 sounds on the basis of the largest cross-product, taking into account the probabilities determined by the preceding phoneme. The recognizer types out the phonemes it recognizes, including spaces. It does fairly well with *t*, *k*, *l*, *m*, *n*, *s*, *sh*, *i*, *a*, *u*, and *uh*, but was not able to recognize *p*, *f*, *th*, or *r*. It was found that, while the use of *a posteriori* probabilities was generally quite helpful, it sometimes resulted in the loss of large parts of words. The machine's performance dropped 50 per cent when it listened to a speaker different from the one for whom it had been adjusted.

Thursday: Meaning and the Human Senses

Thursday's session began with "The Place of 'Meaning' 1 the Theory of Information," by Dr. D. M. Mackay of Kings College, London. Mackay defined the "meaning" f a message in terms of the change it produces in the ecipient's "conditional probability matrix" (array of robabilities of anticipated events). In response to a uestion about "aesthetic information" and the "meaning" f music, Mackay indicated he was willing to include the ffects on the recipient's internal secretions, etc. Bar-Hillel stated that he had a hard time getting anything ut of Mackay's talk and felt it was pointless to define meaning" in terms of even vaguer matrices. A. S. C. Loss also questioned Mackay's ideas.

At this point the Russian delegation finally arrived at ne symposium. There were three fairly young men, vidently all from the Post-Telephone-Telegraph Laboratory in Moscow, and a girl whose job was translator. ubsequent investigation over beer mugs revealed that ney were cordial, but their evident unfamiliarity with poken Western languages prevented any real communication with them.

Professor V. Somenzi of the Istituto di Fisica, Rome, nen delivered his talk "Can Induction be Mechanized?" t amounted to the claim that if the methods of induction an be described precisely, then they can be programmed or an automatic machine. Somenzi does not distinguish etween extrapolation and induction, asserting it is only ecessary to decide which of several possible laws of xtrapolation fits best. In the discussion period, Mackay laimed his own prior invention of an inducing machine, nd someone else pointed out that it has long been known hat any precisely describable operation can in principle e mechanized.

Next, P. Marcou and J. Daguet of the French Centre National d'Études des Télécommunications discussed New Methods of Speech Transmission". The method nvolves feeding the speech wave into a single-sideband uppressed-carrier modulator, passing the output through limiter, and feeding this constant-amplitude signal to a requency divider. In this way, it is possible to compress ne bandwidth of the speech by a factor of as much as 28 when the divider divides by this number. The speech ave is effectively recovered by multiplying the frequency ack to its original value and synchronously demodulating . The resulting output is distinctly different in form from ne original speech wave, on account of the limiter, and known as "constant-level speech," but its spectrum esembles that of the original speech wave, and it sounds ery much like the original, although no explanation has een found for the latter phenomenon. Similar results ave been obtained by digitalizing the system and replacing the speech wave by pulses occurring at its zero crossings.

This constant-level speech makes possible the use of a ass-C rf system, and it yields superior intelligibility in the resence of noise. However, the reduction in redundancy fected by bandwidth compression renders the system susceptible to bad distortion as a result of slight maladjustments. Marcou and Daguet regard the original speech wave as $s(t) = a(t) \cos \phi(t)$, where $a(t)$ is its "instantaneous amplitude" and $d\phi/dt$ is its "instantaneous frequency," which may in fact be either positive or negative. A 90-degree phase shift of all of the Fourier components of $s(t)$ yields its Hilbert transform $S(t) = a(t) \sin \phi(t)$, which would make possible the separation of $s(t)$ into $a(t)$ and $\phi(t)$, $\cos \phi(t)$ being the output of the synchronous detector in the proposed scheme. Unfortunately, the Hilbert transform of $s(t)$ depends upon *all* values of $s(t)$, past, present, and future; hence, such a decomposition of $s(t)$ is not possible while its future is unknown, except perhaps when it is quasi-sinusoidal. Thus, it is not clear what the system does nor how it works, but the fact that it does work is certainly noteworthy. In the discussion period, perhaps as a result, a number of people expressed confusion as to the nature of the system. J. C. R. Licklider of M.I.T. reported that he had listened to the system and that it indeed sounded very good—much better than ordinary clipped speech. E. C. Cherry described this paper as representing good experimental work but lacking in theory; the system preserves the intelligibility of speech, but we don't even know, he said, what it is about speech that it preserves.

D. J. Bruce of the University of Reading Psychology Department then spoke on "The Effects of Context on the Intelligibility of Heard Speech" as measured in the presence of thermal noise. By "context" he meant the general nature of the words in an unconnected list, such as "things to eat." He found that guessing the context correctly is helpful, but guessing it incorrectly can result in an extremely low score.

Next, Professor J. C. R. Licklider of M.I.T. spoke on "Auditory Frequency Analysis." He discussed the incompatability of the place and frequency theories of pitch, and he discussed Huggins's demonstration of the appearance of a pitch when white noise is fed to both ears, arriving at one ear through an all-pass filter; the pitch that is heard is related to a frequency characteristic of the filter, although neither ear alone detects any pitch. This phenomenon requires phase or time analysis and cannot be explained by strict place theory; *i.e.*, by a theory which postulates that every pitch is associated with a particular "place" in the cochlea. By means of a tape recording, Licklider demonstrated that the low pitch associated with a set of three or so of its neighboring higher harmonics can be detected even in the presence of low-pitched thermal noise that completely masks an equally loud fundamental. He concludes that the ears perform running analyses of three types: (1) spectrum, (2) autocorrelogram, and (3) cross-correlogram, and that through a process of association of nerve paths, a person comes to associate the agreeing reports of the three analyses concerning pitch. As a result, one has the sensation of a pitch when the *envelope* of a higher-frequency complex wave has the periodicity associated with a sine wave that pitch. However, if the phase relations of the harmonics are altered, the sensation

of a low pitch may disappear because of the change in the shape of the envelope. Experiments have eliminated the possibility that the low pitch results from nonlinearity and beats. Licklider's theory appears to offer the first unified explanation of a number of auditory phenomena. Afterward, Schouten also discussed and demonstrated the phenomenon of the missing fundamental, using three sine waves between 2 and 3 kc, separated by 200 cps and 200 cps to produce the sensation of 200 cps.

J. T. Allanson and I. C. Whitfield of the University of Birmingham then presented their paper on "The Cochlear Nucleus and its Relation to Theories of Hearing." On the basis of information concerning the inputs and outputs of the cochlear nucleus and available data about its structure, they attempted to determine the type of structure it must have in order to be able to transform the input nervous activity into the output.

Thursday's session ended with a talk by R. L. Gregory of the Cambridge University Department of Experimental Psychology on "An Experimental Treatment of Vision as an Information Source and Noisy Channel." He found an empirical extension of Weber's law $\Delta I/I = C$, where I is the level of illumination and ΔI is the smallest distinguishable increment, to include the effect of the areas A_1 and A_2 of the contrasting fields, namely,

$$\frac{\Delta I}{I} = C_1 + C_2\sqrt{\frac{1}{A_1} + \frac{1}{A_2}}.$$

He regards the breakdown of Weber's law at low intensities as due to "noise" and finds that the law may be extended to cover low intensities by rewriting it in the form $\Delta I/(I + k) = C$. The constant k lies between 0.03 and 0.04 foot-lamberts, but its relationship to the "noise" is not clear. Gregory also attempted to find a function describing the observed relation between area and threshold intensity on the basis of a threshold defined in terms of the probability of mistaking noise for a signal, etc. In the discussion period, Licklider pointed out that Tanner at the University of Michigan asserts there is no sensory threshold—merely a point below which the probability of detection is low. Licklider reported that if an image is stabilized on the retina, visual acuity is momentarily improved, but the image then disappears on account of fatigue. Small movements of the eye therefore do not appear to be responsible for the acuity.

Friday: Behavior and Its Mechanism

J. T. Allanson of the University of the Birmingham University Electrical Engineering Department began the last day's session with a talk on "Some Properties of a Randomly Connected Neural Network." In order to permit a mathematical discussion of the behavior of such a network, he postulated a much simplified model, and derived expressions for the equilibrium rate of firing of neurons, the rate of damping of oscillations, etc., in a fixed randomly connected network, attempting to relate the results to various neural phenomena. Because many rather different networks exhibit very similar gross behavior, Allanson concluded that electroencephalography is not likely to yield detailed information about the structure of the central nervous system.

Next, W. K. Taylor of the University College Department of Anatomy, London, talked on "The Electrical Simulation of Some Nervous System Functional Activities." He described a number of networks built up of model neurons and discussed the functional significance of their stimulus-response characteristics. Examples were included in which the correct response is "learned by association" rather than built into the network; a device was simulated which learns to turn right or left accordingly as rewards are found on the right or left. He concluded that the nervous system can be imitated by a network of interconnected units, most of which receive and deliver pulse-frequency-modulated signals at a multiplicity of input and output terminals; other more specialized units handle the inputs and outputs of the network as a whole. Because the over-all input-output characteristics of the units are nonlinear and depend on past as well as present inputs, it is possible for the network to recognize selected features in visual or auditory patterns and to modify spontaneously generated sequences of activity in a manner which simulates adaptation or learning.

After the morning coffee interval, P. D. Wall delivered a paper by himself, J. Y. Lettvin, W. S. McCulloch, and W. H. Pitts, all of M.I.T., on "Factors Limiting the Maximum Impulse-Transmitting Ability of an Afferent System of Nerve Fibers." Experiments were conducted 1) to determine the maximum sustained frequency of nerve impulses which can be carried along a set of fibers originating in the skin, muscles, and leg joints, proceeding into the spinal column, and streaming toward the head in the dorsal columns of the spinal cord; 2) to determine the limits on impulse transmission following a single impulse or short burst of impulses where all the impulses are carried by the same set of nerve fibers; and 3) to determine the limits following activity in a neighboring parallel set of nerve fibers. The object of the investigation is eventually to calculate the informational capacity of such a system and to ascertain how close to it the rate of transmission of information in the nervous system comes. If it is close, information theory should provide insights for neurophysiology. The quantity of greatest interest is the mean frequency of impulses required to achieve the channel capacity, which is to be compared with the actual average frequency of impulses.

After lunch, Dr. Oliver G. Selfridge of the Lincoln Laboratory, M.I.T., presented a very interesting paper "Pattern Recognition and Learning." The type of learning to which he referred is the evolution of a method for distinguishing patterns, which first requires a motivation and some sort of criterion (*i.e.*, elementary method) for pattern recognition. In particular, he discussed the mechanical recognition of a quadrilateral as it might appear on a snowy television screen. It may be distinguished from a triangle or circle by smoothing to get rid of spots, next eliminating everything but the boundaries between light and dark areas, then eliminating everything but corners or

he ends of edges, replacing each corner by a uniform blob, nd finally counting the blobs. Such experiments in data rocessing have been carried out on an automatic digital omputer. It is Selfridge's feeling that, while a machine an be programmed to learn, there is little hope that a nachine will ever learn to play tolerable chess from the ules alone. In the discussion period, W. K. Taylor laimed a speed for his own analog equipment a million imes as great as that of Selfridge's digital computer.

Next D. E. Broadbent of the Applied Psychology Research Unit, Cambridge, England, under the title "The Concept of Capacity and the Theory of Behavior," resented a plea for the *qualitative* use of information heory in psychology, "to provide a language for talking bout events within the man's nervous system." In the iscussion period, the question of "semantic hygiene" was aised in connection with this proposed use of exact terms n new and vague ways.

After the final tea interval, Henry Quastler of the Jniversity of Illinois Control Systems Laboratory spoke n "Studies of Human Channel Capacity." He has found n experiments with reading a single line of random piano nusic involving an alphabet of 3 to 65 keys in which all otes are equally long, a maximum information rate of 3 bits per second is achieved with a 25-key alphabet. No key is allowed to follow immediately after itself.) Physical limitations reduce the information rate when the lphabet size exceeds about fifty keys or the required peed exceeds about 5.2 keys per second. Quastler found hat the addition of more lines, accompaniment, words, r varied rhythm to the music did not permit a higher nformation rate. This is to be contrasted with Licklider's iscovery of a total rate of 35 bits per second for combined eading and pointing which separately can achieve rates f 25 and 15 bits per second, respectively. Quastler has ot been able to find two such compatible activities. The asks with which high rates are achieved, of course, are hose for which no thinking is required; the information ate in typing runs up to about 17 bits per second.

Similar limitations seem to apply tasks of many types. Quastler concludes that the maximum amount of information of a single kind that can be handled at one time is .5 to 5 bits, the maximum number of kinds of information hat can be handled together is about 7, and the maximum mount of information of all kinds that can be handled at glance is about 20 bits; the maximum information rate r sustained single activities is evidently about 25 bits er second. These limits do not seem to be due to muscular mitations, since proofreading; *e.g.*, which requires little uscular activity, goes at about 20 bits per second. Practice can double the apparent information rate in piano aying though it requires twice the muscular activity.

The last paper on the program was "Application of ommunication Theory to the Human Operator," by D. North of Boulton Paul Aircraft Ltd., Wolverampton, who discussed the applicability of game theory nd of information theory to the study of the characteristics of a human being in a control system.

He was followed by a five-minute talk in English by one of the Russians present, expressing his appreciation for the symposium and the bringing together of people from different countries and sciences; he also expressed his appreciation for Shannon's work. It was his feeling that information theory will yield new high-speed means of communication. He reported that among those in the Soviet Union working in the field of information theory were: Kolmogorov; Khinchin; Yagvo (?); Patelnikov, who works on the theory of potential stability; Kharkevich, who is working on the accommodation of signals and the extension of Shannon's work (he is reported to have published recently in Radiotechnika); Sardoniko, who works on automatic regulation; and the speaker himself, who works on the detection of weak pulsed signals and on active noisy networks. Finally, he expressed his appreciation for the invitation to the symposium sent to his country and for the attention of the audience.[1]

Reprints of some additional papers were made available at the symposium and, though they were not on the program, they may be made part of the record. One of these is "A 'Neuronic Model' of the Inner Ear," by Jørgen Koefoed of Universitetets Fysisk-Kemiske Institut, Copenhagen. Another was "Some Calculations on a Method for Signal Testing," by J. W. Cohen and M. M. Jung of Philips Telecommunication, Hilversum, the Netherlands. Their method is to regard a signal as present if the signal plus noise exceeds some criterion signal over a certain interval, and they have approximated the probability of error of this method.

A count of papers presented at the symposium shows 44 per cent came from Great Britain and 25 per cent from the U. S. Most of the latter came from M.I.T., which appears to be the foremost center of work in the field of information theory.

Publication of Proceedings

The proceedings of the 1950 symposium were published by the IRE, with the cooperation of the Ministry of Supply, in February, 1953. The Butterworths Scientific Press published the 1952 proceedings in September, 1953; and they are to publish the 1955 proceedings shortly. As before, the papers will appear in their formal form rather than as presented, and the remarks made during the discussions periods will be included.

The large number of very stimulating papers presented during the symposium should serve to nourish the field of information theory until the next symposium is held, but perhaps the most important accomplishment of these symposia has been the bringing together of scientists from different fields who are using and extending information theory: physical scientists, philologists, psychologists, neurophysiologists, and mathematicians. The resulting cross-fertilization should prove to be a considerable stimulus to the further development of information theory and its applications.

[1] E. C. Cherry, "Information Theory: Third London Symposium," *Nature*, p. 773; October 22, 1955, identifies the speaker as V. Siforov.

Some Optimal Signals for Time Measurement*

HERBERT SHERMAN†

Summary—Some optimal transmitted signals are given for communications systems in which the time of occurrence of a signal is desired in the presence of additive Gaussian noise. These signals are of two classes; those in which bipolar signal excursions are permitted and those in which only unipolar signal excursions are permissible.

Some bipolar codes used by R. H. Barker are supplemented when conditions permit unlimited negative excursions of the optimally correlated signal output. Other constraints, such as limiting the positive and negative excursions of the optimally correlated signal output, except at the proper phase, will lead to different codes.

When only unipolar codes are permitted, optimum repetitive and nonrepetitive codes (embedded in zeros) are given for various code lengths. The minimum length of such codes is given. A mathematical resemblance is indicated to a frequency allocation problem in which third-order intermodulation products must be avoided.

The closely related concepts of error-detecting and correcting codes and optimally correlated signals are illustrated in the derivation of these codes. The problem of generating such codes by other than trial and error is not solved and remains as a provocative problem.

Introduction

IN CERTAIN communication systems, the time of occurrence or phase of the signal is the parameter which is sought in the presence of additive Gaussian noise. Some examples of such systems are radar ranging, certain navigational systems, PPM systems for voice communications and radar azimuth determination. The ideal signal for such systems are signals which are orthogonal for each possible time position. Such orthogonal signals do not overlap in time. If the received signals are permitted to occupy all possible time positions, the optimal transmitted signal is an impulse function. If the received signal is permitted to occupy only quantized time positions, then the optimal transmitted signal is time limited to the permitted time positions.

Some Optimal Signals for Time Measurement

In many practical cases orthogonal signals are not always possible due to physical constraints. In the radar ranging case, for example, peak power limitations may preclude the short pulse transmission consistent with desired accuracy and available average power. Is there a signal configuration which will make best use of the available energy and also give the best possible phase determination?

In linear systems it has been established by previous writers[1] that the pre-detection cross-correlating filter or "matched" filter (suitably modified for Gaussian colored noise) takes advantage of all of the available energy to create the best peak signal to rms noise ratio in the presence of Gaussian additive noise. Thus, independently of how the signal is shaped in time (providing it is of finite length), we can recover all of the energy into a signal peak, subject only to the difficulties of physical construction. Such a filter might take several forms. One form would be a network whose impulsive response is the reverse in time of the desired signal. In the presence of white noise, the filtered output peak signal squared to average noise power ratio will be $2E/W_0$ where E is the total signal energy and W_0 is the "white" noise power per cps at the filter input.

The peak signal from the cross-correlating filter will indicate (except for time delay through the filter) the time phase of this signal in the absence of noise. Noise will tend to introduce false peaks or to detract from the signal peak. It should therefore be the object of an optimum signal-filter combination to maximize the amplitude separation between the output signal peak and the "side lobes" or leading and trailing edges of the filtered signal output in the absence of noise.

Consider, for example, a pulse-modulated carrier wave which is sampled at twice the carrier frequency and denoted in quantized form

$$\cdots\ 0\ 0\ +1\ -1\ +1\ -1\ +1\ -1\ 0\ 0\ \cdots .$$

A typical output of the optimum filter for this signal would be

$$\cdots 0\,0\ -1\ +2\ -3\ +4\ -5\ +6\ -5\ +4\ -3\ +2\ -1\,0\,0 \cdots .$$

The addition of two units of noise at the (+4) point would create an ambiguity with the (+6) peak signal, while more noise would give a false indication of phase by one or more cycles of the carrier frequency. It is possible to find better signal shapes than this, but some preliminary remarks are in order first.

All of the codes proposed presume that the signal can be given a "fine" structure. In the case of a radar pulse this fine structure may be, for example, 1) frequency or phase modulation within the pulse envelope, or, 2) the division of a long pulse (to increase the average power) into a number of shorter pulses having the same total energy.

It is assumed that a peak power limitation exists and that this power is used fully. There appears to be no remarkable advantage in varying the signal intensity below the peak power level.

The length of the signal may be set by several considerations. The minimum length of the signal may be set by the time necessary to transmit the available energy under constraints of a specified peak power. The maximum

* The major portion of this paper is Chapter VI of a thesis titled "The Detection and Phase Determination of Signals in Additive and Multiplicative Noise," submitted in partial fulfillment of the requirements for the degree of Doctor of Electrical Engineering at the Polytechnic Institute of Brooklyn.

† Lincoln Lab., Mass. Inst. Tech., Lexington, Mass.

[1] See, for example, B. Dwork, "Detection of a pulse superimposed on fluctuation noise," Proc IRE, vol. 38, p. 771; July, 1950.

length of signal may be set by problems of resolution between overlapping signals or by the time needed to give the signal the desired character necessary to optimize the cross-correlated output or by the maximum repetition rate if the signal is quasi-periodic.[2] The time necessary to give the signal the desired characteristics is discussed in the generation of the optimum codes.

Two general types of codes are considered. One may be used where bipolar signals are possible,[3] as in radar systems prior to envelope detection. The second may be used where only unipolar signals are available as is the case after envelope detection or where the signal may be generated serially from pulsed phase-incoherent sources. Ideal bipolar codes have been considered by R. H. Barker.[3] Unipolar codes have arisen in certain frequency assignment problems and have been considered by Babcock[4] and Reiffen.[5]

Both codes attempt to place the bulk of the signal energy in as wide and flat a power spectrum as possible by narrowing the width of the varying portion of the signal autocorrelation function.

Bipolar Codes for Phase Error Correction

In the absence of noise, the output $\phi(\tau)$ of an ideal white noise resistant cross-correlating filter whose desired input is $S(t)$ will be

$$\phi(\tau) = \int_{-\infty}^{0} S(t)S(t-\tau)\,dt.$$

In the absence of signal and noise, the output of the filter will naturally be zero. It is apparent then that the maximum amplitude separation we can usefully achieve between signal and no-signal or between signal and phase-shifted replica, is that between the output signal extreme and zero signal level. We would like this to be achieved at some interval (which will also condition the occupied frequency spectrum) consistent with our desired smallest interval of phase determination. This can be expressed as an integral inequality.

$$\int_{-\infty}^{0} S(t)S(t-\tau)\,dt \le 0; \qquad |\,\tau\,| > \tau_1.$$

The negative excursions of the auto-correlation function $[\phi(t)]$ will be important where multiple signals are present since they may combine with the noise to subtract from a signal pulse and cause detection failures. For this purpose it would be desirable but not always possible to require that the output signal be identically zero for all $|\,\tau\,| > \tau_1$. This may be inconsistent with the time required to transmit the desired amount of energy with limited peak power. Some negative excursion will therefore be found necessary.

The integral inequality can be reduced to the following form for substantially band-limited (W) signals by sampling $S(t)$ at $S(t_1) = S_1$, $S(t_2) = S_2, \cdots$, where $S(t) = 0; t > t_N; t_2 - t_1 = 1/2W$ and considering values of τ which are integral multiples of the sampling period

$$\begin{aligned}
&S_N S_1 \le 0\\
&S_N S_2 + S_{N-1}S_1 \le 0\\
&S_N S_3 + S_{N-1}S_2 + S_{N-2}S_1 \le 0\\
&\quad\vdots\\
&S_N S_{N-1} + S_{N-1}S_{N-2} + \cdots + S_2 S_1 \le 0.
\end{aligned}$$

There are N unknowns to be found from $N - 1$ equations with $N - 1$ adjustable constants (on the right-hand side). It is helpful and physically acceptable to restrict S_i to values of $+1$, -1, and zero (a ternary code).

Solutions exist for all values of $N \ge 2$. This existence theorem is readily demonstrated by allowing

$$S_N = +1$$
$$S_1 = -1.$$

These values separated by a string of zeros will always satisfy the inequalities given above. This is a trivial solution however. At least two extremes are interesting. In the first it would be desirable to have every pulse position occupied (by other than zero). This will permit the maximum peak in the auto-correlation function in the shortest total signal time interval. R. H. Barker[3] has sought these solutions and found them for $N = 3, 7$, and 11.

Failing this, one would like to maximize the distance between output signal peaks and "side lobes" of the correlated signal output. This implies that one should consider "side lobes" in the correlated signal output other than zero and, in fact, should make the maximum off-peak signal output a dependent variable. Limitations of manual computation precluded exploration of some of these byways.

Another extreme among the possible solutions is to require

$$\begin{aligned}
&S_N S_1 = -1\\
&S_N S_2 + S_{N-1}S_1 = 0\\
&\quad\vdots\\
&S_N S_{N-1} + S_{N-1}S_{N-2} + \cdots + S_2 S_1 = 0.
\end{aligned}$$

[2] An assumption is made here that Doppler frequency shifts generated by motion of the transmitter and receiver are small (*e.g.*, the period of the Doppler frequency shift is long compared to the signal length). Woodward and Davies discuss uncertainty relations due to Doppler shifts. In addition, while this paper was in manuscript form, unpublished work by Siebert and others at Lincoln Laboratory, M.I.T., concerning the uncertainty effects of Doppler shift and on ternary codes similar to those described hereafter came to the attention of the author.

[3] R. H. Barker, "Group synchronization of binary digital systems," "Communication Theory," W. Jackson, Ed., Academic Press, New York, N. Y., p. 273; 1953.

[4] W. C. Babcock, "Intermodulation interference in radio systems," *Bell Sys. Tech. Jour.*, p. 63; January, 1953.

[5] B. Reiffen, "A System for Minimizing Intermodulation Interference in Communication Systems Based on Frequency Selection," Lincoln Lab. Group Rep. 311–5, July, 1955.

TABLE I
"OPTIMUM" TERNARY CODES

$N = 2$	− +		
*$N = 3$	− + +	− 0 +	
$N = 4$	− 0 0 +		
$N = 5$	− 0 − + +	− 0 + 0 +	
$N = 6$	− + − 0 + +	− 0 0 0 0 +	
*$N = 7$	− + − − + + +	− 0 0 + 0 0 +	
$N = 8$	− + − − 0 0 + +		
$N = 9$	− + − 0 0 − 0 + +	− 0 0 0 − 0 + 0 +	− 0 0 0 − 0 0 0 +
*$N = 11$	− + − − + − − − + + +		

+, − are pulses of unit amplitude and given sign. Mirror images are not shown.

* The codes having no zeros are Barker's.

TABLE II
AUTOCORRELATION FUNCTIONS FOR OPTIMUM TERNARY CODES

$N = 2$	− +2		
$N = 3$	− 0 +3	− 0 +2	
$N = 4$	− 0 0 +2		
$N = 5$	− − 0 0 +4	− 0 0 0 +3	
$N = 6$	− 0 0 0 − +5	− 0 0 0 0 +2	
$N = 7$	− 0 − 0 − 0 +7	− 0 0 0 0 0 +3	
$N = 8$	− 0 0 −2 0 0 0 +6		
$N = 9$	− 0 0 0 − 0 0 − +6	− 0 − 0 0 0 0 0 +4	− 0 0 0 0 0 0 0 +3
$N = 11$	− 0 − 0 − 0 − 0 − 0 +11		

This approximates the condition one might prefer where the "apparent" resolution of super-imposed signals is a problem.

The solution of the array of inequalities is somewhat of a chore. The number of possible solutions is 3^N. For $N = 9$, this implies 19,683 possibilities. Trial and error should be avoided if possible.

While some consistent procedures for solving these inequalities were evolved, these involved a good deal of trial and error and led to solutions for optimum ternary codes of the former type, given in Table I along with Barker's codes. The autocorrelation functions for these codes are shown in Table II. The auto-correlation functions are symmetric about the maximum. Only half of the autocorrelation function is therefore shown.

BINARY ERROR-CORRECTING PHASE-DETECTING CODES

As an approach to the determination of an optimum signal for phase-detection where only unipolar signals are available, the concept of a *binary* error-detecting and correcting code similar in function to the previously discussed ternary code will be used. In the case of certain multiplicative noise, these unipolar binary codes become most appropriate for phase discrimination.[6]

Ideally it would be desirable to find a binary code sequence which would, in the presence of digit errors, unerringly indicate the correct phase of the signal. Unfortunately one can only approximate such a sequence by selecting the sequence which will correct for one or more errors. We therefore seek a binary code sequence which will unambiguously indicate the phase of a signal in the presence of errors of less than a given maximum number.

The phases of a cyclic signal sequence are given by one of all possible sequences generated by advancing each signal simultaneously by one pulse position and feeding the "carry" digit into the start of a sequence. As an example, all possible phases ϕ of one sequence are:

ϕ_1	1 1 0 0 0
ϕ_2	0 1 1 0 0
ϕ_3	0 0 1 1 0
ϕ_4	0 0 0 1 1
ϕ_5	1 0 0 0 1

It will be assumed hereafter that each signal may occupy all possible phases. The number of phases which can occur is the *sequence length.* A noncyclic group of digits is hereafter called a message. Its phase is advanced by one digit movement of the entire message.

For the purpose of accounting for error we shall use the concept of "distance" between codes. The distance between two binary sequences is equal to the minimum number of digit changes required to convert one sequence into the other. The concept of a minimum distance as a measure of correct signal choice is identical to choice of the maximum cross-correlated signal in the case of phase discrimination. The use of distance in this section happens to be convenient for the derivation of optimum *binary* codes. The relationship between the usual definition of squared distance (hereafter called "distance")

[6] H. Sherman, *loc. cit.*

$$d^2(\tau) = \sum_i [S_i(t) - S_i(t + \tau)]^2,$$

and the number of digit changes can be demonstrated for binary codes by making the physical subtraction required by $[S_i(\tau_1) - S_i(\tau_2)]$. When the digits are identical this term goes to zero; when the digits are unlike, the differences are ± 1, which when squared and added leads to the definition given above.

If we expand the bracket and combine terms

$$d^2 = 2[\sum_i S_i^2 - \sum_i S_i(t)S(t + \tau)],$$

we can recognize "distance" as twice the difference between the peak of the autocorrelation function and the value of the autocorrelation function at τ.

The maximum number of errors which can be tolerated is less than half of the minimum "distance" between phases. If the number of errors exceed this, the received message may be ambigious (exactly half-way between two possible messages) or erroneous (closer to the wrong message).

One possible statement of the problem is to fix the number of 1 (signal) digits (similar in certain cases to fixing the signal energy), which are here assumed to be less than the zero digits, and to determine the sequence length and digit sequence in which the maximum number of digit errors can be corrected. For this message the "distance" between all possible phases of the message shall be equal to or greater than a given number.

The problem is nontrivial only in the case where successive signal phases overlap (*i.e.*, the phase difference between signals is less than the message length where the *message length* refers to the smallest number of digits which encompasses all 1 digits in any phase). When the signal phases *do not* overlap, then the "distance" between signal phases is always equal to twice the number of 1 digits in the message and is independent of the disposition in 1 digits and the length of the message.

Thus ϕ_1 and ϕ_2 are two signal phases (nonoverlapping) which are always four units apart in either of the following cases

ϕ_1 1 1 0 0 0 0
ϕ_2 0 0 0 1 1 0

or

ϕ_1 1 0 1 0 0 0
ϕ_2 0 0 0 1 0 1

For overlapping, phase-shifted binary sequences, the "distance" between least distant phases can be no greater than twice the number of 1 *digits (assuming these are less in number than zero digits) less* 2.

The greatest "distance" between two identical (except in phase) binary sequences is twice the number of 1 digits, if the 1 digits are fewer in number than 0 digits. The signal, while phase shifted through all possible phases, must have *at least* one position in which a digit coincidence occurs between phases. In general the distance between two phases is twice the number of 1 digits less twice the number of 1 coincidences between phases. If a minimum of one digit coincidence occurs, then the distance between closest phases is as stated above.

Let us examine an example of an error-corrected code.

Example Number of signal digits = 3
Sequence length = 7
Maximum digit error tolerated = 1
Minimum distance between phases = 4

1 1 0 1 0 0 0
0 1 1 0 1 0 0
0 0 1 1 0 1 0
0 0 0 1 1 0 1
1 0 0 0 1 1 0
0 1 0 0 0 1 1
1 0 1 0 0 0 1

The generation of such codes may be accomplished by a trial and error procedure.[7]

It will be useful to characterize the successful code by a series of decimal digits which denote the digit positions from a 1 digit to the next succeeding 1 digit. If the sequence were cyclic as in

1 1 0 1 0 0 0 1 1 0 1 0 0 0 1 1 0 1

then the decimal sequence would be

1-2-4-1-2-4-1-2

This particular sequence will therefore be denoted as a 1-2-4 sequence (or 4-1-2 or 2-4-1 or inverting, 4-2-1, 1-4-2, 2-1-4). *It is important to note that this sequence, as well as all succeeding optimum codes,*

1) has no repetition of decimal digits within one cycle,
2) is such that the sum of any two or more contiguous decimal digits is unequal to any other digit in the sequence or to the sum of any other two or more contiguous decimal digits.

This condition is necessary since its violation would permit *two* digit positions to coincide in two phases reducing the minimum distance by an additional 2 digits from the optimum.

The example given above was for an optimum error-correcting code of minimum length for a cyclic repetition of the sequence. As a result of the requirement that the sum of contiguous decimal digits (characterizing the spacing of 1 digits) shall be unequal to any other sum of

[7] The reviewer has pointed out a paper presented to the 1955 Third London Symposium on Information Theory by D. A. Huffman, "The Synthesis of Linear Sequential Coding Networks," which presents methods for finding shorter error-correcting codes (known therein as "maximum length null sequences") in which the ones and zeros are nearly equal. Such codes are useful in the important cases where signal energy is dependent only on the length of the code rather than on the number of one digits.

TABLE III
BINARY ERROR-CORRECTING CODES

Max. Errors Corrected	Min. Distance	No. of 1 digits	Min. Cyclic Sequence Length	Decimal Code	Binary Code	Min. Message Length	Decimal Code	Binary Code
0	2	1	2	0	0 1	1	0	1
0	2	2	3	1-2	1 1 0	2	1	1 1
1	4	3	7	1-2-4	1 1 0 1 0 0 0	4	1-2	1 1 0 1
2	6	4	13	1-3-2-7	1 1 0 0 1 0 1 0 0 0 0 0 0	7	1-3-2	1 1 0 0 1 0 1
3	8	5	21	1-3-10-2-5	1 1 0 0 1 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0	12	1-3-5-2	1 1 0 0 1 0 0 0 0 1 0 1
4	10	6	31	1-3-2-7-8-10		18	1-3-6-2-5	

contiguous decimal digits we can establish the minimum length of such codes. Since each sum must be different from all others, the number of sums which can be formed from an N digit sequence length is

1) the decimal digits taken singly $= N$
2) the decimal digits taken in contiguous pairs $= N$
3) the decimal digits taken in triplets $= N$

.
.
.

4) the decimal digits taken in "$(N - 1)$ lets" $= N$
5) the decimal digits taken N at a time $= 1$

The minimum number of sums is therefore

$$L = N(N - 1) + 1.$$

There can be no fewer sums than this or two would be equal. Since the largest sum of contiguous decimal digits is equal to the sum of all decimal digits (the length of the sequence), this largest sum must equal or exceed L in order to have room for the minimum number of sums, L.

The length of the sequence must therefore be greater than or equal to L.

If the sequence can be embedded in a number of zeros such that the sequence length exceeds the minimum above, then we can group our 1 digits to have shorter message lengths. Some examples of both are given in Table III.

The mathematical problems of finding the optimum binary error-correcting code for phase determination, where the code is embedded in zeros, is identical to that of specifying the allocation of frequencies which are free of third order intermodulation products. This problem has been considered by two authors. Mr. W. Babcock[4] has found the most dense frequency allocations (error-correcting codes in this context) for up to 10 frequencies (digits). These are reproduced in Table IV. Mr. B. Reiffen[5] has found sets (not the most compact sets) of frequency allocations for up to 33 frequencies (digits). The signal digits in Mr. Reiffen's set are located in the following positions, between which fall zeros.

TABLE IV
MOST COMPACT CODES UP TO 10 DIGITS*
EMBEDDED IN A SEQUENCE OF ZEROS

The number indicates the location of 1 digit.

0, 1, 3
0, 1, 4, 6
0, 1, 4, 9, 11
0, 1, 4, 10, 12, 17
0, 1, 4, 10, 18, 23, 25
0, 1, 4, 9, 15, 22, 32, 34
0, 1, 4, 13, 24, 30, 38, 40, 45
0, 1, 7, 11, 26, 39, 47, 56, 59, 61

* W. E. Babcock, *loc. cit.*

0	30	147	400	821	1416
1	44	181	474	915	1522
3	65	203	564	969	1650
7	80	251	592	1015	
12	96	289	661	1158	
20	122	360	774	1311	

ACKNOWLEDGMENT

I am indebted to Professor A. E. Laemmel of the Polytechnic Institute of Brooklyn for the suggestion of the error-detecting code as an approach to the optimal signal for phase determination.

An Analysis of Signal Detection and Location by Digital Methods*

G. P. DINNEEN† AND I. S. REED†

Summary—**An analysis of the detection and location of repetitive gnals in noise by digital techniques is made. The problem of locaon of the center of signals, herein denoted as beam-splitting, is xplored. A Monte Carlo method employing a high speed digital omputer was used to obtain quantitative results for a variety of igital detectors. A method of mathematical analysis is described nd used to check computed results. The work described differs om much of the previous literature on detection or statistical desion theory in that an estimate of signal location is demanded.**

Introduction

RECENTLY the detection of repeated signals in noise has been aided by signal-integration schemes using digital methods for signal storage.[1,2] These hemes have generally employed a method of integration hereby quantized signals are added. Although it is ossible to quantize the signal amplitude into many disrete levels, we shall consider quantization into two mplitude levels and between fixed time markers. If the riginal signal plus noise waveform exceeds a predeermined amplitude between given time markers, a stanard pulse is generated at the end of the interval; if the reshold is not exceeded, no pulse is generated. The robability of obtaining a standard pulse can then be etermined and the probability of the sum of pulses over fixed interval is given by the binomial distribution. The dvantage of these digital methods over a corresponding nalog scheme is the possibility of remembering signals ver a large number of samples. Thus the increase in gnal-to-noise ratio as the square root of the number of mples can be fully realized. As the number of samples is creased however, one suffers a loss in accuracy, since we ow know only that a signal is present somewhere within e sample interval. Moreover, if the signal straddles two tervals we have an added ambiguity. In order to overome these disadvantages, a scheme for locating the nter of the signal in addition to its detection is called for.

For purposes of exposition, the quantizer can be condered to deliver a one if the signal is above the quantizer reshold, and a zero if below. Thus the quantizer may e regarded as transforming a radar sweep into a sequence ones and zeros, where each digit is delivered to the detector at the quantizer time intervals. Imagine now that each quantized sweep sequence is laid parallel in time sequence from left to right in such a manner that the information at a given range is along a horizontal line. Time measured in sweep number would be the abscissa, and range would be the ordinate of this picture (Fig. 1). A target would appear as an increased density of ones at a given range throughout an interval of sweeps corresponding to the beamwidth of the radar.

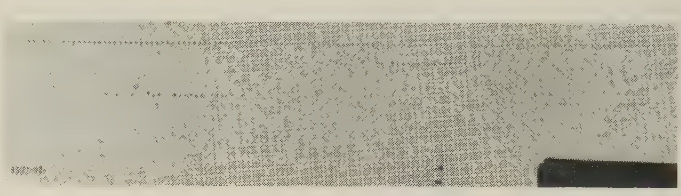


Fig. 1—Quantized radar video.

If a machine has enough memory to hold the target information for several sweep times, on the order of a beamwidth, then an estimate of the density of ones can be made. If this estimate of density exceeds some threshold, one can say that a target has been detected. The number of false targets detected by this means can be kept low by a proper setting of the quantizer threshold.

If changes in density, as well as the density, can be discovered, an estimate of the azimuthal positions of the leading and trailing edges of a target sequence at some range can be obtained. By this information, it is clear that an azimuthal estimate of the center of the target is obtained. In fact, if $\Delta\theta$ is the azimuth increment between leading and trailing edges of the target, and if θ is the azimuth of the trailing edge, then $\theta - \Delta\theta/2$ is an estimate of the target center. The general process discussed above has been termed the beam-splitting process.

It is the purpose of this paper to discuss various methods for detection and location of signals in noise. This represents a natural extension of the digital integration scheme discussed in Harrington's paper.[2]

Signal-to-Noise Ratio

For purposes of exposition, let us consider the threshold detection of a sine wave in additive Gaussian noise. The probability density for the envelope R of a sine wave plus noise (P sine $\omega t + I_n$), where the noise is confined to a relatively narrow band centered on the sine-wave frequency, is

$$p(R) = \frac{Rd\,R}{\psi_0} \exp\,[-(R^2 + P^2)/2\psi_0] I_0\left(\frac{RP}{\psi_0}\right), \qquad (1)$$

* The research in this document was supported jointly by the my, Navy, and Air Force under contract with the Massachusetts stitute of Technology.

† Mass. Inst. Tech., Lincoln Laboratory, Lexington, Mass.

[1] J. V. Harrington and T. F. Rogers, "Signal-to-noise improvement rough integration in a storage tube," Proc. IRE, vol. 38, pp. 97–1203; October, 1950.

[2] J. V. Harrington, "An analysis of the detection of repeated gnals in noise by binary integration," Trans. IRE, vol. IT-1, . 1–9; March, 1955.

where ψ_0 is the noise power and I_0 is a modified Bessel function of the first kind and zero order.[3] For quantization in time of the order of a pulsewidth, the probability of obtaining a quantized video pulse, given a pulse of amplitude P in the IF amplifier, is

$$P_s = P(a, v) = \int_V^\infty V\, d\, V \exp\left[-(V^2 + a^2)/2\right] I_0(aV),$$

where the variables have been normalized with respect to the noise power; *i.e.*, $V = R/\sqrt{\psi_0}$, the predetector signal-to-noise ratio $a = R/\sqrt{\psi_0}$, and V is the normalized amplitude threshold in the quantizer.

In the special case where a signal is not present ($a = 0$), the probability of obtaining a quantized video pulse due to noise alone is

$$P_n = P(0, V)$$
$$= \int_V^\infty V\, d\, V \exp\left[-V^2/2\right] = \exp\left[-V^2/2\right]. \qquad (2)$$

Using these two equations, the functional relationship among a, P_s, and P_n may be deduced. This is shown graphically in Fig. 2.

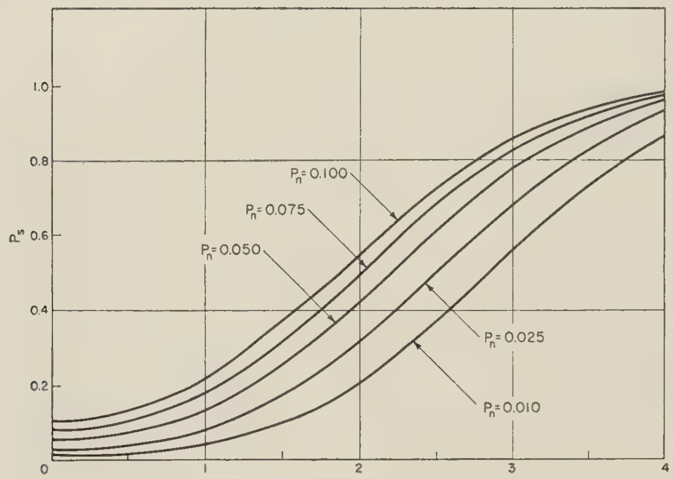


Fig. 2—P_s = probability of obtaining a quantized video pulse due to signal, P_n = probability of obtaining a quantized video pulse due to noise alone. a = predetector signal-to-noise ratio.

Hence, in later discussions, when reference is made to predetector signal-to-noise ratio, the a used in the equation above is implied. This model has been used to provide some measure of signal strength for various settings of P_s and P_n.

Target Description

Let us first remark that the quantization into zero or one, of course, destroys all amplitude information except whether or not the signal exceeds a certain threshold. Therefore, precise information about the beam pattern is not useful unless a real symmetry is apparent; that is, unless the density of ones over some interval less than the beamwidth increases as one approaches the center of the target and decreases on the far side. In many cases, the fading and scintillation causes enough gaps in the target so that this symmetry does not occur. Therefore, it is assumed that the primary indication of the presence of a target is a sharp increase in the real density of ones in *some* interval corresponding to the beamwidth of the radar antenna.

Mathematical Model

A mathematical model of the target and noise regions after quantization can be given in terms of Bernoulli sequences. A Bernoulli sequence is a sequence of trials, each of which has two possible outcomes with probabilities p and $q = 1 - p$. For convenience, the outcomes will be denoted by one with probability p and zero with probability $q = 1 - p$. A target region is then a subsequence of a Bernoulli sequence with $p = P_s$, and a noise region is then a subsequence of a Bernoulli sequence with $p = P_n$.

Consider one of the quantized sweep sequences corresponding to a given range as in Fig. 1. Azimuthal positions will be measured now in sweep numbers. The target that appears as an increased density of ones has a width in sweep numbers corresponding to the beamwidth of the radar antenna. The mathematical model for one of these fixed range sweep sequences is a pair of Bernoulli sequences. The quantized radar video, that is, the zero-one sequence, corresponding to the noise region, is represented by a Bernoulli sequence with $p = P_n$ and the zero-one sequence, corresponding to the target region, is represented by a Bernoulli sequence with $p = P_s$.

This description of the target region emphasizes its statistical nature. For example, if the average number of hits per beamwidth is 16 and $p = P_s = 0.31$, then Fig. 3 is a plot of the probability distribution of real densities. This is, of course, just the binomial distribution. Crossed points are those obtained by means of the stochastic process carried out on Whirlwind I, described later.

Blip-Scan Ratio

The blip-scan ratio is the ratio of visible returns from a target to the number of scans where blip is used to denote visible return. For a human observer and a scope this means the number of times the observer sees the target divided by the number of scans or the number of times when observation is possible. The integration over a beamwidth is then performed by the scope and by the eye of the observer. Experimental data are usually taken for a particular radar to plot the dependence of the blip-scan ratio against signal strength or signal-to-noise ratio. This curve will be a function of the radar and of the observer.

[3] S. O. Rice, "Mathematical analysis of random noise," *Bell Syst. Tech. Jour.*, vol. 23, pp. 282–332; July, 1944; vol. 24, pp. 46–156; January, 1945.

For this study, the integration over a beamwidth is erformed by the detector. The final blip-scan ratio is gain dependent on the radar receiver and also on the haracteristics of the detector. The value of P_s which we hall use, defined as the probability of a one due to signal lone, may be analogously called the average hit-sweep atio. P_s may be considered as the mean of the ratio of he number of ones returned from a target to the total umber of sweeps per beamwidth. Fig. 3 shows the distri-ution of the hit-sweep ratio for an assumed value of $P_s = 0.31$; that is, out of 1,000 targets, about 210 will ave a hit-sweep ratio of 5/16. The blip-scan ratio for a etector will be the probability of detection, or Q, as it vill be denoted later. The curves that will be drawn later vill give the blip-scan ratio as a function of P_s, that is, he average hit-sweep ratio. It is interesting to note that f one plots the blip-scan ratio vs the actual (not average) it-sweep ratio, the graph has sharp discrimination and mphasizes the essential nature of all the detectors tudied, which is that of measuring real density over some rbitrary interval.

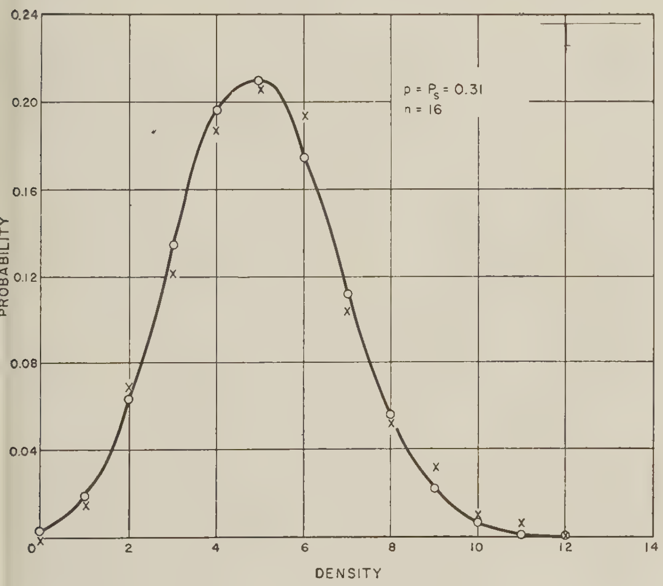


Fig. 3

The Detectors

Since a target region is distinguished from a noise egion primarily by its density, the design of a digital letector for quantized video is equivalent to the problem of measuring the density over some interval determined by the beamwidth. The detector must also be physically ealizable, which means primarily that its storage require-nents are not excessive.

This problem is much more difficult than would appear rom the above formulation. The detector, of course, will ave no *a priori* knowledge of the target beginning. Therefore, it must be sufficiently sensitive that it detects he increased density region quickly, but not so sensitive that it detects false targets due to noise. If it does detect a target beginning, the detector must perceive the end of the increased density region. However, if it is too sensitive to this change, it will have a tendency to split targets. The detector design criteria can be stated as follows.

1) It must detect the maximum number of targets with the minimum number of false alarms.

2) It must have quick decision time on target beginning, but must not exceed a prescribed number of false decisions.

3) It must have quick release time in noise, but must not exceed a prescribed number of split target decisions.

It is clear that the design of a detector must be a compromise because of the conflict in the design criteria and the constraint of small memory.

Although it is possible to make a theoretical analysis of many special detectors, much remains to be done along these lines. In particular, it would be desirable to have empirical relations between such quantities as beam-splitting accuracy and detection sensitivity, beam-splitting accuracy and storage capacity, etc., which would apply to large classes of detectors.

For this paper, three classes of detectors have been chosen on the basis of previous theoretical and experimental work. These detectors and their modifications will be described in the following sections.

Success-Run Detector

It is possible to calculate the probability of occurrence of a specific run; *i.e.*, configurations of zeros and ones, in a Bernoulli sequence of zeros and ones.[4] The mean recurrence time, or the time between happenings of these runs, may also be calculated. Certain runs have a high probability of occurrence, hence a short mean recurrence time in a noise region; and a low probability of occurrence, hence a long mean recurrence time in a signal region. Thus, a simple success-run observer would count the recurrence time of some event and make a decision between target and noise based on the magnitude of this count. This would be done most simply by setting a threshold and, if the count exceeds the threshold, the observer detects a target, otherwise noise. In contrast with Feller's studies,[4] overlapping runs will be considered. Thus, if the event is 0 0 0, it is said to occur 4 times in the following sequence

$$\begin{array}{cccccccccc} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ & & \uparrow & \uparrow & \uparrow & & & & \uparrow \end{array}$$

For example, consider the following, where S measures basic time units (sweeps in the radar case, V is the result of the Bernoulli trial (quantized video in the radar case) and L is the counter which measures recurrence times.

[4] W. Feller, "An Introduction to Probability Theory and Its Applications," John Wiley and Sons, Inc., New York, N. Y., 1950.

S	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
V	0	0	0	0	0	1	0	1	1	0	1	1	1	0	1	1	0	0	1	0	0	0	0	1	0		
L	0	0	0	0	0	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	0				

In this case the L-counter, or recurrence-time counter, reaches 16 before it is reset to zero by the occurrence of the event 0 0 0. On the other hand, if the Bernoulli sequence has a small probability $p = P_n$,

S	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
V	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	
L	0	0	0	1	2	3	4	5	6	0	0	1	2	3	0	0	0	0	1	2	3	0

The largest count is 6. Thus, experimentally one might set a threshold at 7 so that, if the counter does not reach 7, the decision is noise, but if it exceeds 7, the decision is target. To find the center of the Bernoulli sequence one takes $L/2$ and subtracts from S.

The modified success-run observers detect on the basis of the recurrence of one of several events. The assumption is made that the observer is trying to detect a target in noise so that the L-counter is set at zero and remains at zero until a one is received. It then looks for the occurrence of one or more events E_1 throughout an interval L_1. If E_1 occurs, the counter is reset to zero and the process begins again. If E_1 does not occur, the detector searches for an event or events E_2 through an interval L_2. If E_2 occurs, the counter is reset to zero and the process begins again. If E_2 does not occur, the detector searches for an event or events E_3 through an interval L_3, etc. Based on an *a priori* probability of the length of the target region, a threshold T is set. Whenever the counter is reset to zero, the value l_n of the L-counter just prior to resetting is compared with T. If $l_n \leq T$, it is assumed that the count is due to noise and no action is taken. If $l_n \geq T$, it is assumed that the count is due to signal, and $l_n/2$ is subtracted from the value of the S-counter, or sweep counter, to give the location of the center of the target sequence. Different detectors are obtained by choice of the intervals $L_1, L_2, L_3, \cdots$, the events $E_1, E_2, \cdots$, and the threshold T.

A special case will now be described, but first the operation of one of these modified detectors will be contrasted with a simple success-run detector (Fig. 4).

Success Run A. The L-counter of Success Run A is initially set at zero. When a one is received, the L-counter is set to one at the next unit (sweep). The L-counter obeys the following rules. At each unit time (sweep), an observation of $\mu_{n-3}\mu_{n-2}\mu_{n-1}\mu_n$ and l_n is made, where μ_n is the value of the Bernoulli trial at the nth sweep, and l_n is the value of the L-counter at the nth sweep. If

1) $0 \leq l_n \leq 3$ and either $\mu_{n-3}\mu_{n-2}\mu_{n-1}\mu_n = 0\ 1\ 0\ 0$

or $\mu_{n-2}\mu_{n-1}\mu_n = 0\ 0\ 0$

2) $3 < l_n \leq 6$ and $\mu_{n-2}\mu_{n-1}\mu_n = 0\ 0\ 0$

3) $6 < l_n$ and $\mu_{n-3}\mu_{n-2}\mu_{n-1}\mu_n = 0\ 0\ 0\ 0,$

then

$$l_{n+1} = 0; \quad \text{otherwise} \quad l_{n+1} = l_n + 1.$$

A target is detected when l_n exceeds 6. The target beginning occurs for $l_n = 1$ and the end occurs when $l_n > 6$ and $l_{n+1} = 0$. There is, therefore, a natural bias of 2.5.

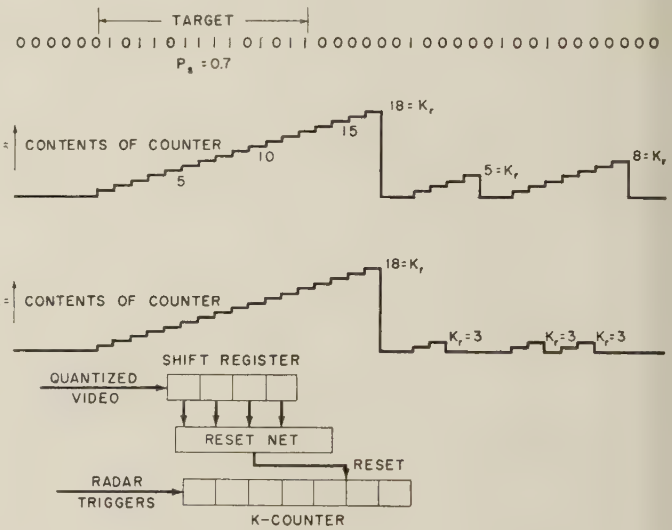


Fig. 4—QV = Quantized video in given range interval. Success-run detector reset when 4 successive zeros occur in QV. Modified success-run detector reset when: 1) Last 3 QV pulses are 000 or 100 and $l < 3$, 2) Last 3 QV pulses are 000 and $4 < e < 7$, 3) Last 4 QV pulses are 0000 and $7 < l < 63$.

Success Run B. At each unit time (sweep), an observation of $\mu_{n-4}\mu_{n-3}\mu_{n-2}\mu_{n-1}\mu_n$ and l_n is made. If

1) $0 \leq l_n \leq 4$ and either $\mu_{n-3}\mu_{n-2}\mu_{n-1}\mu_n = 0\ 1\ 0\ 0$

or $\mu_{n-2}\mu_{n-1}\mu_n = 0\ 0\ 0$

2) $4 < l_n \leq 7$ and $\mu_{n-2}\mu_{n-1}\mu_n = 0\ 0\ 0$

3) $7 < l_n$ and $\mu_{n-4}\mu_{n-3}\mu_{n-2}\mu_{n-1}\mu_n = 0\ 0\ 0\ 0$

then

$$l_{n+1} = 0; \quad \text{otherwise} \quad l_{n+1} = l_n + 1.$$

A target is detected when l_n exceeds 7. The target beginning occurs for $l_n = 1$, and the end occurs when $l_n > 7$ and $l_{n+1} = 0$. There is, therefore, a natural bias of 2.5.

equential-Observer Detectors

The sequential-observer detectors are a class of detectors ıat consist of an up-down counter with an upper threshold ıd a lower threshold. The counter is so designed that it ounts up r steps when a one is received, and down s eps when a zero is received. When the upper threshold reached, the detector senses a target. When the lower ıreshold is reached, the detector senses noise. By varying ıe values of r, s, and the positions on the upper and lower ıreshold, one obtains a large number of different sequenal-observer detectors. The operation of one of these etectors is illustrated in Fig. 5.

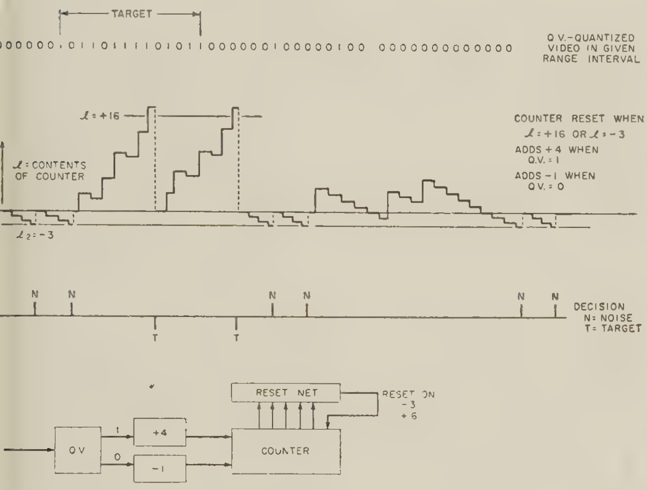


Fig. 5

Sequential Observer A_1. At each sweep time, or unit me, an observation of μ_n is made. If $\mu_n = 1$, the sequential ounter, which is originally set at $+0$, is caused to count p 4; if $\mu_n = 0$, the sequential counter is caused to count own one. If the new value of the counter is ≥ 16, $_{+1} = +0$, and a target pulse is generated. If the new alue of the counter is ≤ -3, then $l_{n+1} = +0$, and a o-target pulse is generated. The signal is measured from ıe first target pulse to the first no-target pulse following it.

Sequential Observer B_1. At each sweep time, or unit time, ı observation of μ_n is made. If $\mu_n = 1$, then $l_{n+1} = l_n + 2$; $\mu_n = 0$, $l_{n+1} = l_n - 1$, with the added provision that, if $_{+1} \geq 7$, l_{n+1} is changed to 2. If $l_{n+1} \leq 0$, l_{n+1} is changed 2. The signal width is measured from the first target ılse, when $l_{n+1} \geq 7$, to the first no-target pulse following , when $l_{n+1} \leq 0$.

'oving Average Detectors

The class of moving average detectors makes decisions ısed on the density of ones inside some interval. The idth of the interval is determined by the *a priori* probility of the width of the target region. The rules for the -counter, or density counter for a width l, would be

$$
\begin{aligned}
d_{n+1} &= d_n + 1 \quad \text{if} \quad \mu_n = 1 \quad \text{and} \quad \mu_{n-l} = 0 \\
d_{n+1} &= d_n - 1 \quad \text{if} \quad \mu_n = 0 \quad \text{and} \quad \mu_{n-l} = 1 \\
d_{n+1} &= d_n \quad \text{if} \quad \mu_n = 0 \quad \text{and} \quad \mu_{n-l} = 0 \\
& \qquad \text{or} \quad \mu_n = 1 \quad \text{and} \quad \mu_{n-l} = 1.
\end{aligned}
$$

If the sum of ones inside this "window" exceeds a predetermined threshold, it is classed as a signal region, otherwise a noise region.

A target beginning is sensed when one passes from noise to signal and a target end is sensed when one passes from signal to noise. The target center can then be calculated by taking the midpoint. By varying the "window" width and the detector threshold setting, a large number of detectors of this class may be formed.

Computer Analysis

A series of programs has been coded for Whirlwind I which simulate radar inputs and the several detectors under consideration. Using this stochastic model, measurements have been made for the purpose of evaluating these detectors.

For purposes of this study, a fixed range has been assumed; *i.e.*, one of the horizontal lines in Fig. 1. The spread of the simulated azimuth pulses is about 9 beamwidths of the radar. In approximately the center of the azimuth spread, a target corresponding to one beamwidth has been simulated.

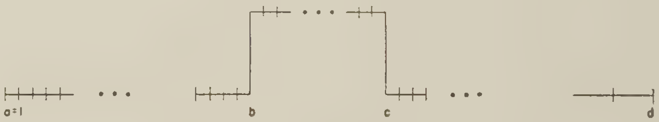


Fig. 6

Fig. 6 is a representation of the simulated input. In the interval $a = 1$ to b, a Bernoulli sequence with $p = P_n$ is generated; in the interval b to $c - 1$, a Bernoulli sequence with $p = P_s$ is generated; and in the interval c to d, a Bernoulli sequence with $p = P_n$ is generated. For almost all runs, $b = 80$, $c = 96$ and $d = 144$. A sequence of 15-bit random numbers is generated, using a pseudo-random number generator. These are then split into two 7-bit random numbers. Each 7-bit random number has a value between 0 and 127. A test number M is chosen, and if the random number is less than or equal to M, a one is generated; if the random number exceeds M, a zero is generated. Thus, if $M = 63$, a Bernoulli sequence is generated with $p = 0.5$. If $M = 3$, a Bernoulli sequence is generated with $p = 0.03$. Thus, in this model the radar quantized video is simulated by the outputs of two Bernoulli sequences—one with $p = P_n$, the other with $p = P_s$. The width of the target region is chosen to agree with the assumed beamwidth.

This composite Bernoulli sequence is repeated a large number of times—for most runs, 1,000 times. The density

in the noise regions and in the target regions is given by the binomial distribution.

The several detectors were simulated by programming Whirlwind I to carry out the actual operations of the detector. Briefly, for each of the 1,000 runs of simulated radar data, the coded detector acting on the basis of the incoming zeros and ones made decisions as to target or noise. When a target was detected, the center was located with reference to the units $a = 1$ to d of the particular run.

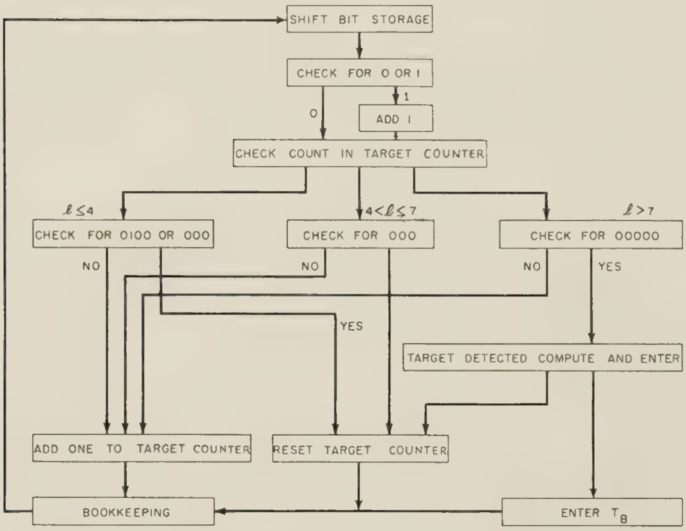


Fig. 7

For example, the flow chart which is a symbolic representation of the operations to be performed by a program for the Success Run B observer is given in Fig. 7. From this flow chart, the coded program for Whirlwind I was written. In a similar way, the other detectors were programmed.

Location	Number of Target Centers Detected
83	1
84	9
85	15
86	17
87	12
88	38
89	40
90	39
91	38
92	32
93	36
94	18
95	4
96	2
97	1

Fig. 8—Results for Success Run B.

For purposes of evaluation, several measurements were recorded and printed out. For each detector and each pair of settings P_s, P_n, the detected target centers were recorded in units $a = 1$ to d. Thus, for example, for Success Run B ($P_n = 0.03125$, $P_s = 0.3125$), the values given in Fig. 8 were printed out. The left-hand column gives the position from $a = 1$ to $d = 144$, and the righ hand column is the number of target centers detected a each location. A similar measurement of target-beginnin locations was made for some detectors. The total of th right-hand column gives the number of targets detecte in 1,000 runs.

Using a desk calculator, the mean and variance of th distribution of target center locations was determined.

It is possible, of course, that a detector may rese inside a target region and thus detect one target as tw targets. Therefore, the number of times no targets, on target, two or more targets were detected was recorded.

For purposes of evaluating the radar input model, th density in the intervals $a \leq x \leq b - 1$, $b \leq x \leq c -$ 1 $c \leq x \leq d - 1$ were measured. Also, the frequency c ones at any given unit was recorded. Several other auxi iary measurements such as these were made to check ou the coded programs.

Data obtained are presented by graphs [Figs. 9 and 1 (opposite) and Figs. 11-13 (p. 36)], one for each detecto and each setting of P_n, P_s. The graphs give the distributio of target-center locations with reference to the center c the generated signal region; *i.e.*, if the target region is 8 $\leq x \leq 95$, the center is taken to be $x = 88$. The followin information is included with each graph.

1) The values of P_s, P_n and a (see Target Description

2) The value of Q (the percentage of targets detectec the probability of target detection, or blip-scan ratio);

3) The value of FA (the probability of false-targe detection);

4) The value of ST (the probability of detecting a singl target as two or more targets);

5) The values of m and σ, (the mean and variance of th distribution of target-center decisions).

Theoretical Calculations

The equations for the probabilities of the various state of the L-counter of the success-run observer were writte out. A detailed study of this kind of analysis has bee made by Reed.[5,6] In order to illustrate the method, th equations for Success Run A will be written out in detai as follows:

$$x^1_{n+1} = p_n x^0_n$$

$$x^2_{n+1} = p_n x^1_n + p_{n-1} q_n x^0_{n-1}$$

$$x^3_{n+1} = p_n x^2_n + p_{n-1} q_n x^1_{n-1}$$

$$x^4_{n+1} = p_n x^3_n + p_{n-1} q_n x^2_{n-1} + p_{n-3} p_{n-2} q_{n-1} x^0_{n-3}$$

$$x^5_{n+1} = p_n x^4_n + p_{n-1} q_n x^3_{n-1} + p_{n-2} q_{n-1} q_n x^2_{n-2}$$

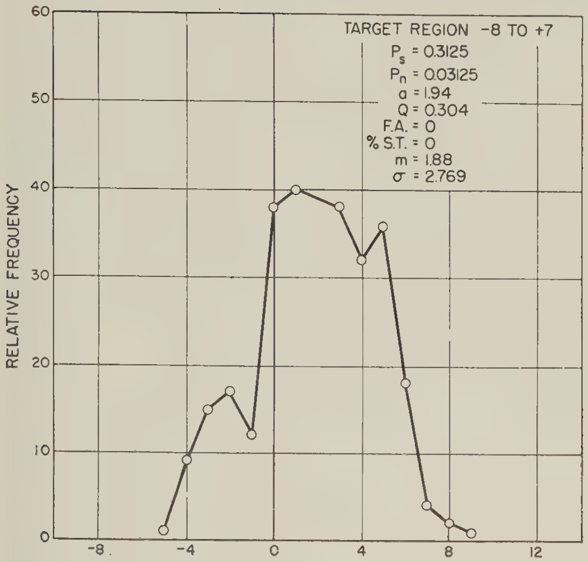
$$x^6_{n+1} = p_n x^5_n + p_{n-1} q_n x^4_{n-1} + p_{n-2} q_{n-1} q_n x^3_{n-2}$$

$$x^7_{n+1} = p_n x^6_n + p_{n-1} q_n x^5_{n-1} + p_{n-2} q_{n-1} q_n x^4_{n-2}$$

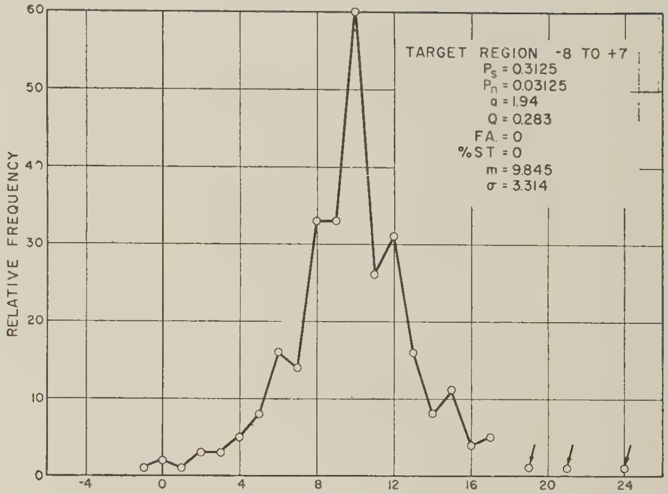
$$x^8_{n+1} = p_n x^7_n + p_{n-1} q_n x^6_{n-1} + p_{n-2} q_{n-1} q_n x^5_{n-2} + p_{n-3} q_{n-2} q_{n-1} q_n x^4_{n-}$$

[5] I. S. Reed, Lincoln Lab. Tech. Rep. (not generally available).

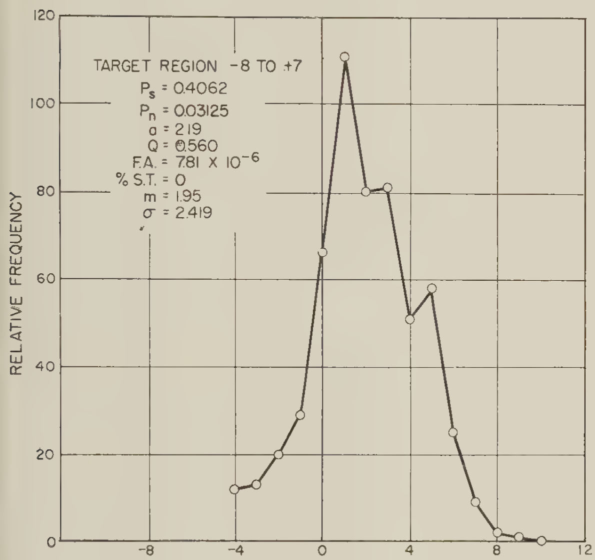
[6] I. S. Reed, Lincoln Lab. Tech. Rep. (not generally available



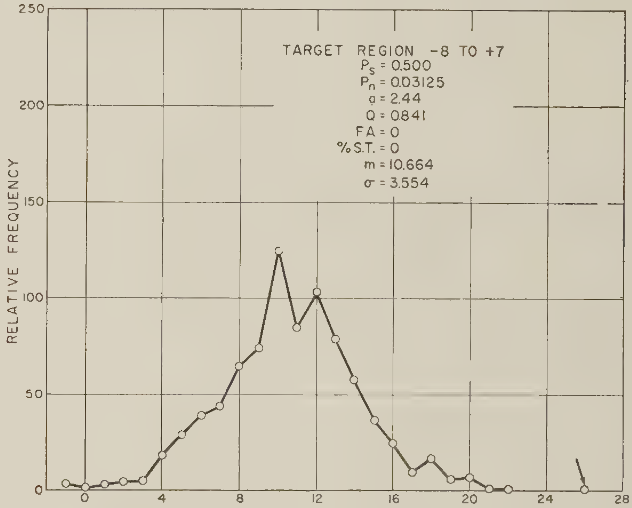
Target Center Position—Success Run *B*.



Target Center Position—Sequential Observer *A*.

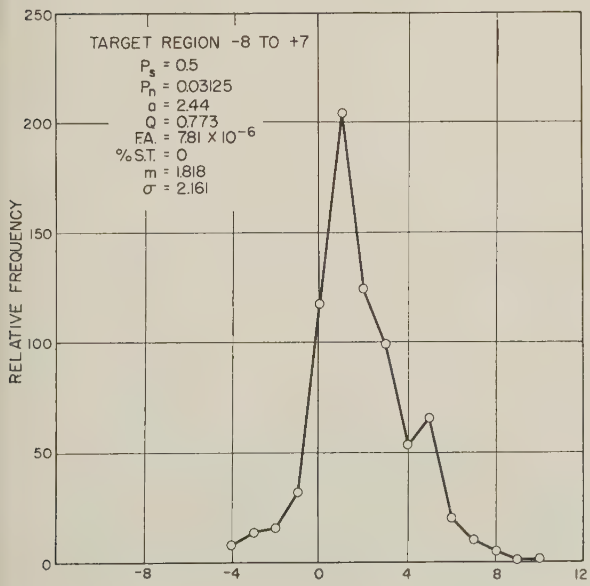


Target Center Position—Success Run *B*.



Target Center Position—Sequential Observer *A*.

Fig. 10



Target Center Position—Success Run *B*.

Fig. 9

$$x_{n+1}^{9} = p_n x_n^{8} + p_{n-1} q_n x_{n-1}^{7} + p_{n-2} q_{n-1} q_n x_{n-2}^{6} + p_{n-3} q_{n-2} q_{n-1} q_n x_{n-3}^{5}$$

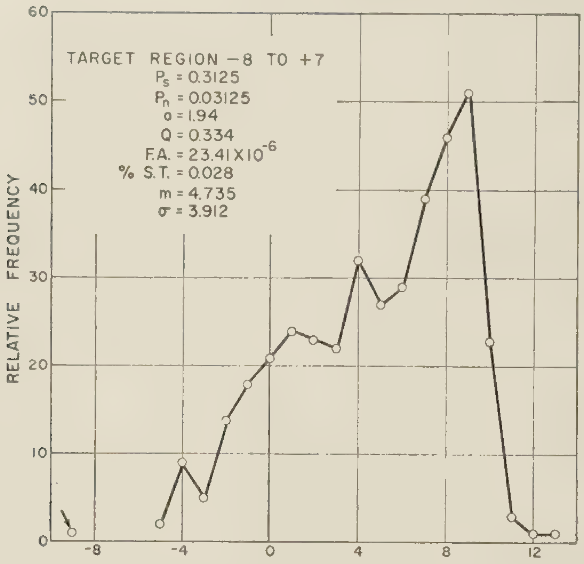
$$x_{n+1} = p_n x_n^{8} + p_{n-1} q_n x_{n-1}^{7} + p_{n-2} q_{n-1},$$

where x_{n+1}^{j} is the probability that the counter has the value j at time $n+1$, and p_n is the probability of a one at time n. These equations are derived by assigning probability measures to the sets of Bernoulli sequences.

In a heuristic manner, the equations are derived by considering the reset conditions and writing down the probability of the other events. For example, to get x_{n+1}^{5}, one knows that the count must have been 4 at time n. The reset conditions for count 4 are 0 0 0. Thus, a one in the first place will always cause the counter to advance (this corresponds to $p_n x_n^{4}$). A zero in the first place and a one in the second place will give x_{n+1}^{5} (this corresponds to $q_n p_{n-1} x_{n-1}^{3}$). A zero in the first two places preceded by a one will give x_{n+1}^{5} (this corresponds to $q_n q_{n-1} p_{n-2} x_{n-2}^{2}$). Thus,

$$x_{n+1}^{5} = p_n x_n^{4} + p_{n-1} q_n x_{n-1}^{3} + p_{n-2} q_{n-1} q_n x_{n-2}^{2}.$$

By similar reasoning, the other equations may be derived.



Target Center Position—Sequential Observer B_1.

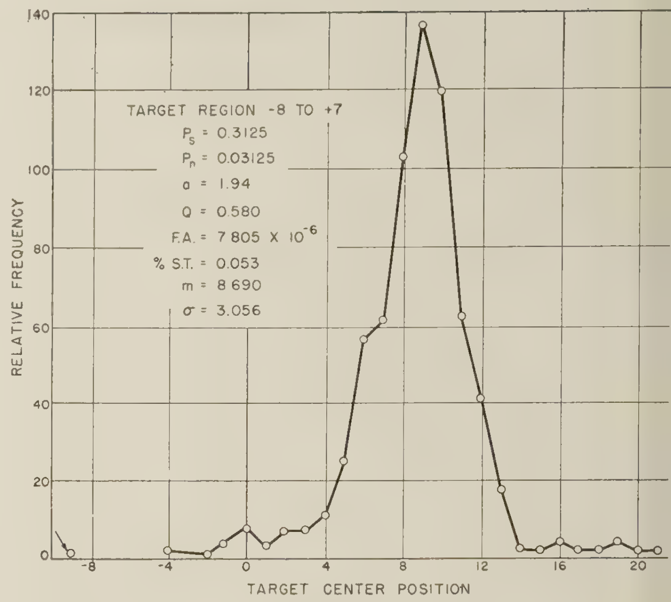
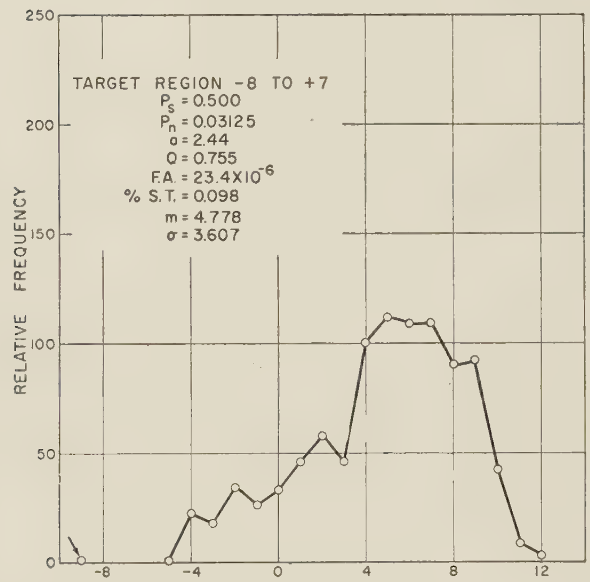


Fig. 12—Moving average detector, window 16, $T = 5$.



Target Center Position—Sequential Observer B_1.

Fig. 11

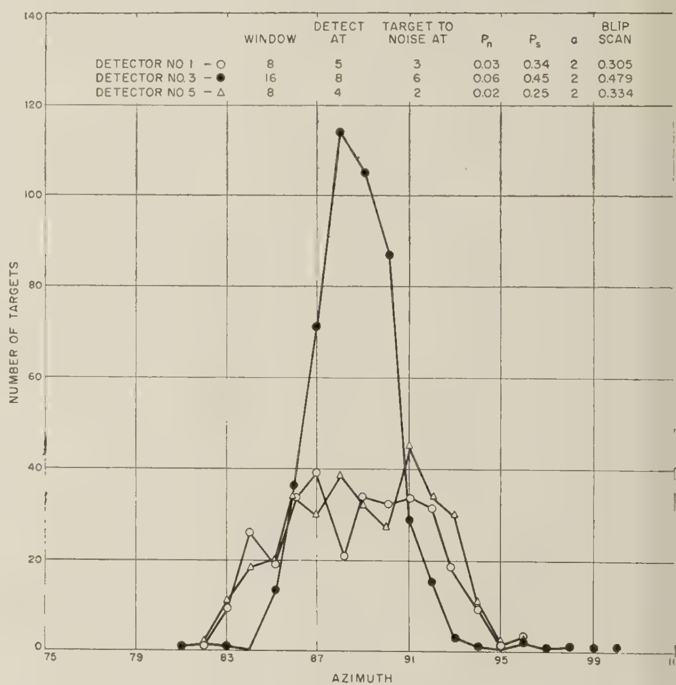


Fig. 13—Moving average detectors.

In order to solve the transient problem, that is, the determination of the probability of detection of a signal region following a long noise region, it is necessary first to calculate the steady-state probabilities in noise, which are then used as the initial conditions for the transient equations.

The steady-state equations are obtained from Success Run A by removing the subscripts n which refer to time, since it is now assumed that these probabilities are independent of time. We then obtain the following steady-state equations for Success Run A.

$$x^1 = px^0$$

$$x^2 = px^1 + pqx^0 = p^2x^0 + pqx^0 = px^0$$

$$x^3 = px^2 + pqx^1 = p^2x^0 + p^2qx^0 = p^2(1 + q)x^0$$

$$x^4 = px^3 + pqx^2 \qquad + p^2q^2x^0$$

$$x^5 = px^4 + pqx^3 + pq^2x^2$$

$$x^6 = px^5 + pqx^4 + pq^2x^3$$

$$x^7 = px^6 + pqx^5 + pq^2x^4$$

$$x^8 = px^7 + pqx^6 + pq^2x^5 + pq^3x^4$$

$$x^9 = px^8 + pqx^7 + pq^2x^6 + pq^3x^5$$

$$x = px^8 + pqx^7.$$

	P_n	P_s	a	Q	F.A.
Success Run A	0	0.3		0.387	
Success Run A	0.03	0.3	1.94	0.424	
Success Run B	0	0.3		0.265	
Success Run B	0.04	0.275	1.7	0.198	24×10^{-6}
Success Run B	0.05	0.310	1.7	0.274	56×10^{-6}
Success Run B	0.06	0.342	1.7	0.357	112×10^{-6}

Fig. 14—Table of calculated values.

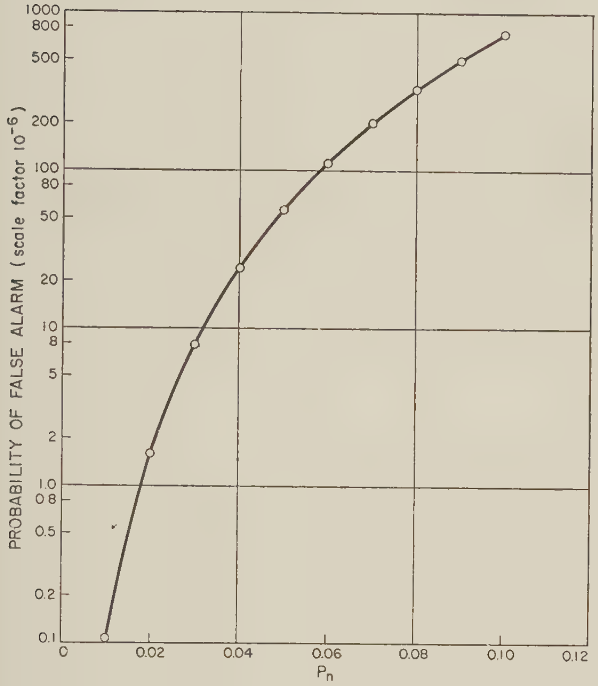


Fig. 15—False alarm rates.

To solve these equations, we sum the columns of this infinite system, using the identity that $\sum_{i=0}^{\infty} x^i = 1$, and obtain

$$1 - x^0 = p + pq + pq^2(1 - x^0 - x^1) + pq^3(1 - x^0 - x^1 - x^2 - x^3) + p^2q^2x^0,$$

or

$$x^0 = \frac{q^4}{1 - pq^2 - pq^3 - 2p^2q^3 - p^3q^3 - p^3q^4}.$$

With the value of x_0 obtained by substituting numerical values of $p = P_n$ and q, the values of x^i for steady-state noise can be obtained. Then the procedure is to calculate x_1^j, $j = 1, 2 \cdots 8$, if $S_n = \sum_{j=7}^{\infty} x_n^j$, or the probability that the counter has a value exceeding 6. Assuming that within a signal region no reset for $l > 6$ occurs, $S_{n+1} = S_n + x_{n+1}^8$. Therefore, $S_1 = S_0 + x_1^7$. Finally, $x_1^0 = 1 - \sum_{j=1}^{6} x_1^j - S_1$. The second column is computed in just the same way. The value of S_{16} is the probability that the count has value exceeding 6 at time $n = 16$ or the probability of detecting the signal. These calculations were carried out using a desk calculator. The results are given in Fig. 14.

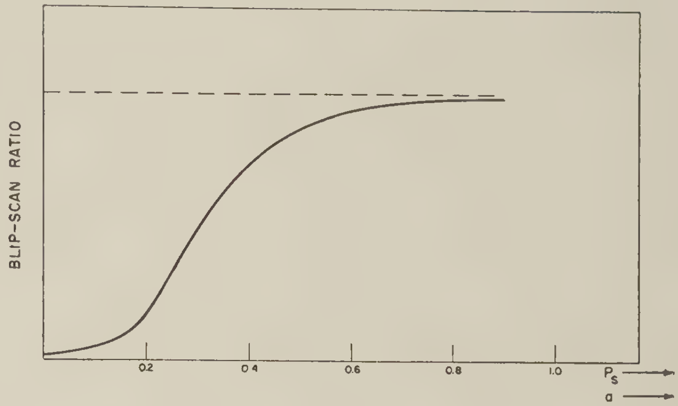


Fig. 16—Target detection curve.

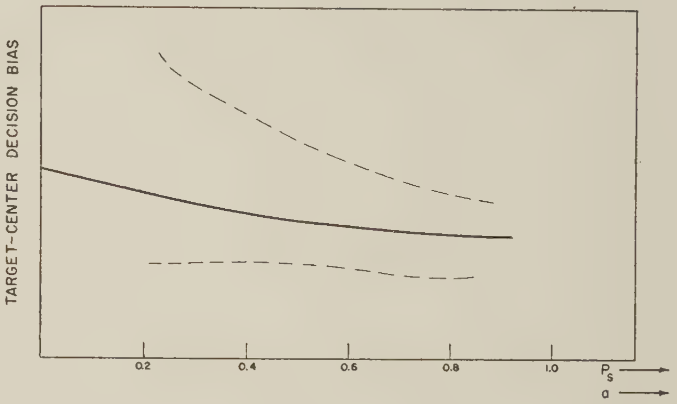


Fig. 17—Target center decisions.

Information about the false-alarm rate can be obtained from the steady-state equations for noise. This was done for various values of P_n and the results collected in Fig. 15, where x_8 is the probability of detecting a false target.

The equations for any other detector in this class may be written in just the same way and similar calculations made. It is possible to program these equations on Whirlwind I.

Conclusion

A good detector should have a high blip-scan ratio for targets and a low false-alarm rate; *i.e.*, sharp discrimination between target and noise. This will be referred to as the detection criterion, and may be represented graphically by a curve of blip-scan ratio vs P_s for a fixed P_n, or a, as in Fig. 16. For good discrimination, the detection curve should have a sharp knee.

A good beam-splitter should have a decision bias which is very nearly independent of signal strength and has a very narrow variance. This will be referred to as the beam-splitting criterion, and may be represented graphically by a curve of decision bias vs P_s for fixed P_n, or a, as in Fig. 17.

A prohibitive amount of calculations would be required to obtain curves for both detection criteria and beam-splitting criteria for each detector. Therefore, it was

necessary to extrapolate from just a few points of these curves. In most cases, results were obtained for two values of P_s (0.3 and 0.5) for $P_n = 0.03$. These results have been shown graphically.

The analysis described here consisted primarily of a computer study using a Monte Carlo Method. Sufficient exact analysis was carried out to compare certain check points. By these methods, any proposed beam-splitter could be analyzed. A very large number of detectors were studied which are not reported here. Furthermore, consideration was given to other methods employing amplitude information or beam shape. The general conclusion is that the moving window detector satisfies the detection and beam-splitting criteria and at the same time is logically the simplest. A moving window detector can provide accuracy of target center with a variance of two or three units, or sweep times. This is so close to the theoretical optimum that it does not seem expedient for most applications to complicate the detector for the very small increment in accuracy. Of course the resolution time of these detectors is limited to a beamwidth.

Although a Monte Carlo method using a high speed digital computer provides information about proposed detectors, the solution of the statistical problem would be more useful in predicting the performance of new detectors and in evaluating the influence of parameters such as beam-width, signal strength, or beam pattern.

Acknowledgment

The helpful advice and criticism received from many associates at the Lincoln Laboratory is gratefully acknowledged. A special acknowledgment is due Miss B. Jensen for carrying out the extensive numerical calculations and for the preparation of many of the figures.

Inverse Probability in Angular-Tracking Radars*

B. J. DUWALDT†

Summary—An angular-tracking radar of the pulsed type is analyzed for targets whose signal-return in the video follows a Rayleigh distribution. The pulse amplitudes are assumed independent from pulse-to-pulse but closely correlated for the duration of a pulse width. White Gaussian noise is introduced into the input of the receiver. It is assumed that rf phase information is lost. The method of inverse probability is used to find the most probable direction of the target from the scan axis. It is found that interference may be reduced by synchronizing transmissions in the scan cycle, by transmitting pulses in pairs simultaneously or nearly simultaneously, and by increasing the number of pulses integrated.

Introduction

THE RADAR considered here is a pulsed radar system in which the rf phase information is lost. The receiver is assumed to consist of linear, band-limiting mixer-IF stages and an ideal linear envelope detector. White Gaussian noise is introduced in the first stage of the receiver.

It will be assumed that the nature of the targets is such that the received pulse amplitudes will have a Rayleigh distribution[1-4] and, further, that the amplitudes are independent from pulse to pulse. The noise out of the detector between pulses will, of course, also have a Rayleigh distribution.[5] When operating in the tracking mode, the radar antenna will scan in the vicinity of the target such that the mean-square signal-to-noise ratio at the detector output will be varied in accordance with the angular position of the target with respect to the scan axis. With a knowledge of the beam-shape and pulse-transmission locus, the probability distribution of target angular position may be derived from observation of the detector output, henceforth designated y. This *a posteriori* probability distribution is found from the relationship

* This paper is based on part of a doctoral thesis submitted August, 1955 to the faculty of Purdue University. The work was supported by the U. S. Naval Ordnance Plant, Indianapolis, Indiana, under Department of the Navy Contract No. Nonr 1100-(03).

† Ramo-Wooldridge Corporation, Los Angeles, California. This work was done while the author was a Research Instructor in the School of Electrical Engineering, Purdue University, Lafayette, Indiana.

[1] J. L. Lawson, and G. E. Uhlenbeck, "Threshold Signals," M.I.T Radiation Laboratory Series, vol. 24, p. 249, McGraw-Hill, New York; 1950.

[2] D. E. Kerr, Ed., "Propagation of Short Radio Waves," M.I.T. Rad. Lab. Ser., vol. 13, McGraw-Hill, New York, Chapt. 6; 1951.

[3] R. H. Delano, "A Theory of Target Glint or Angular Scintillation in Radar Tracking," Convention Record of the IRE, part I, p. 13; 1953.

[4] P. Swerling, "Probability of Detection for Fluctuating Targets," RAND Corp. Res. Memo. RM-1217; March 17, 1954.

[5] S. O. Rice, "Mathematical Analysis of Random Noise," *Selected Papers on Noise and Stochastic Processes*, p. 214, Dover Publications, Inc., New York; 1954. This paper originally appeared in *Bell Sys. Tech. Jour.*, vols. 23 and 24, 1945.

$$\textit{A Posteriori} \text{ Probability Distribution of Position} = \frac{\textit{A Priori} \text{ Probability Distribution of Position} \times \text{Likelihood Function of } y}{\textit{A Priori} \text{ Probability Distribution of } y} \quad (1)$$

where the likelihood function of y is the probability distribution of y given a target position.[6-7]

To use the method of inverse probability, it is first necessary to consider the *a priori* distribution of the pulse amplitudes. The beam power pattern is assumed to yield a mean-square signal-to-noise ratio of

$$a = a_{\max} \exp \frac{-1.4\theta^2}{B^2}, \quad (2)$$

where $a_{\max}$ is the mean-square signal-to-noise ratio when the target is on the beam axis, θ is the angular distance of the target from the beam axis, and B is a measure of beamwidth. This signal-to-noise ratio assumes that pulse transmission and reception occur at the same angle with respect to the beam axis. The signal-to-noise ratio is ζ^2/σ^2 where $2\zeta^2$ is the mean-square value of y when there is a target but no noise and $2\sigma^2$ is the mean-square value of y when noise alone is present.

The antenna beam axis during transmission and reception always lies on the surface of a cone whose axis is the scan axis. The angle between the scan axis and beam axis is θ_0. The target lies on a sphere having its center at the radar antenna. The surface area on this sphere circumscribed by the locus of the transmission beam axis may be approximated, for small θ_0, by a plane surface perpendicular to the scan axis and through the target. The angles between the beam axis, scan axis, and target line-of-sight will then be approximately directly proportional to the line segments joining the intersections of these axes on the plane and the target as long as the angles are small (say less than 20°). The plane is shown in Fig. 1 in which R is the range of the target from the radar antenna.

From the geometry of the figure,

$$\theta^2 = \theta_0^2 + \Delta^2 - 2\theta_0\Delta \cos(\psi_n - \phi), \quad (3)$$

where Δ is the angular distance of the scan axis from the target, ψ_n is the angular rotation of the beam axis from the reference line at time t_n (the time of transmission), and ϕ is the angle of the target from the reference line. From (3),

$$a(\Delta, \phi, \psi_n) = a_{\max} \exp \frac{-1.4}{B^2} [\theta_0^2 + \Delta^2 - 2\theta_0\Delta \cos(\psi_n - \phi)]. \quad (4)$$

The signal-to-noise ratio a_0 when the target is on the scan axis is

$$a_0 = a_{\max} \exp \frac{-1.4\theta_0^2}{B^2}. \quad (5)$$

Hence, (4) may be written

$$a(\Delta, \phi, \psi_n) = a_0 \exp\left[\frac{-1.4\Delta^2}{B^2} + \frac{2.8\theta_0\Delta}{B^2} \cos(\psi_n - \phi)\right]. \quad (6)$$

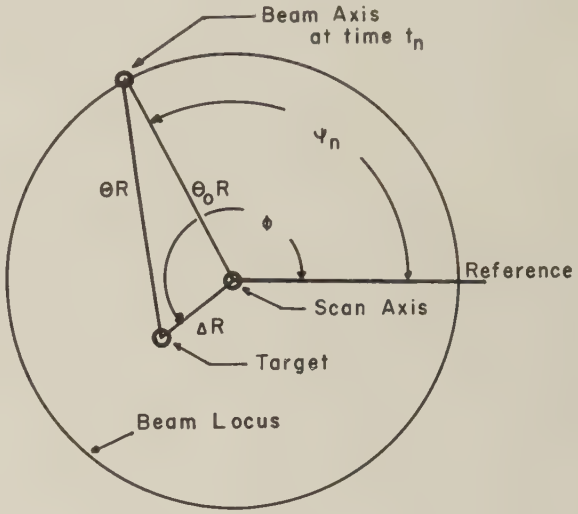


Fig. 1—Plane approximation of target area on the tracking sphere.

A Priori Probability Distribution of Pulse Amplitude

One difficulty that arises in analyzing a tracking system is that the antenna is part of a closed loop. The amount the scan axis is likely to be off the target depends upon how well deviations of the scan axis from the line-of-sight are detected and how well the control system returns the scan axis to the line-of-sight. Hence, the *a priori* probability distribution of pulse amplitude will be a function of the error constants of the servo system in the tracking loop. For a first-order analysis, however, it may be assumed that the radar does a good job in keeping Δ very small. It is reasonable to assume, then, that the *a priori* probability distribution of pulse amplitude will be a Rayleigh distribution resulting from a mean-square signal-to-noise ratio a_0.[8] If k pulses are considered, let $\{y_n\} = (y_1, y_2, \cdots, y_k)$ represent the received signal from a set of k pulses. Since there is no dependence between pulses, the joint *a priori* probability of having this set of k pulses is simply the product of the probability distributions of the individual pulses. Hence, for Rayleigh-distributed pulses, the *a priori* distribution becomes

$$p[\{y_n\}] = \frac{\prod_{n=1}^{n=k} y_n}{[\sigma^2(1 + a_0)]^k} \exp\left[\sum_{n=1}^{n=k} \frac{-y_n^2}{2\sigma^2(1 + a_0)}\right]. \quad (7)$$

[6] P. M. Woodward, and I. L. Davies, "Information theory and inverse probability in telecommunications," *Proc. I.E.E.*, vol. 99, part III, p. 37; March, 1952.
[7] P. M. Woodward, "Probability and Information Theory, with Applications to Radar," McGraw-Hill, New York; 1953.

[8] The *a priori* probability distribution of pulse amplitude for some particular parameters is presented in the Appendix of the thesis cited.

Likelihood Function and a Posteriori Probability Distribution of Target Angular Position

In a like manner, the likelihood function of $\{y_n\}$ given Δ and ϕ, is

$$p_{\Delta,\phi}[\{y_n\}] = \frac{\prod_{n=1}^{n=k} y_n}{\prod_{n=1}^{n=k} \sigma^2[1 + a(\Delta, \phi, \psi_n)]} \cdot \exp\left[\sum_{n=1}^{n=k} \frac{-y_n^2}{2\sigma^2[1 + a(\Delta, \phi, \psi_n)]}\right]. \tag{8}$$

The *a priori* joint probability distribution of (Δ, ϕ) depends upon the nature of the targets and the radar system tracking loop. It seems likely that with land-based radar tracking aircraft targets there will be more error in bearing than in position angle. On the other hand, a radar system on a rolling ship may have more error in position angle than in bearing. Hence, the most likely assumption, in general, is a flat distribution for ϕ and a Rayleigh distribution for Δ.

The *a posteriori* joint probability distribution of (Δ, ϕ) is, substituting (7) and (8) into (1),

$$p_{\{y_n\}}(\Delta, \phi) = p(\Delta, \phi) \frac{(1 + a_0)^k}{\prod_{n=1}^{n=k} [1 + a(\Delta, \phi, \psi_n)]} \cdot \exp\left[\sum_{n=1}^{n=k} \frac{y_n^2[a(\Delta, \phi, \psi_n) - a_0]}{2\sigma^2(1 + a_0)[1 + a(\Delta, \phi, \psi_n)]}\right]. \tag{9}$$

If the *a priori* distribution of Δ is such that Δ is almost always very much smaller than B, then $a(\Delta, \phi, \psi_n)$ may be approximated by

$$a(\Delta, \phi, \psi_n) = a_0\left[1 + \frac{2.8\theta_0\Delta}{B^2} \cos(\psi_n - \phi)\right]. \tag{10}$$

If a further simplification is made by using a_0 for $a(\Delta, \phi, \psi_n)$ except where the difference of these two quantities occurs, then the *a posteriori* distribution becomes

$$p_{\{y_n\}}(\Delta, \phi) = p(\Delta, \phi) \exp\left[\sum_{n=1}^{n=k} \frac{y_n^2 a_0}{2\sigma^2(1 + a_0)^2} \cdot \frac{2.8\theta_0\Delta}{B^2} \cos(\psi_n - \phi)\right]. \tag{11}$$

Decision on Direction of the Target from the Scan Axis

If there is no reason to weight the decision on ϕ, the most probable value may be used. The value of ϕ which maximizes the exponential is found by differentiation. The exponential is maximum when its argument is maximum. Ignoring the multiplier which is independent of n and ϕ, the function to be maximized is

$$\begin{aligned} H &= \sum_{n=1}^{n=k} y_n^2 \cos(\psi_n - \phi) \\ &= \sum_{n=1}^{n=k} y_n^2 \cos \psi_n \cos \phi + \sum_{n=1}^{n=k} y_n^2 \sin \psi_n \sin \phi \\ &= X \cos \phi + Y \sin \phi, \end{aligned} \tag{12}$$

where

$$X = \sum_{n=1}^{n=k} y_n^2 \cos \psi_n \qquad Y = \sum_{n=1}^{n=k} y_n^2 \sin \psi_n . \tag{13}$$

Differentiating with respect to ϕ, and equating to zero,

$$\frac{dH}{d\phi} = Y \cos \phi - X \sin \phi = 0$$

gives a maximum when

$$\cos \phi = \frac{X}{\sqrt{X^2 + Y^2}} \qquad \sin \phi = \frac{Y}{\sqrt{X^2 + Y^2}}, \tag{14}$$

where the square root is positive. The ϕ specified by (14) is, therefore, the most probable value.

Quality of the Decision

Assuming that (14) is the best decision on ϕ, the quality of the decision depends upon parameters such as pulsing pattern, beamwidth, scan angle, and number of pulses. The first part of the criterion of quality to be used here will be that the average values of X and Y simultaneously become directly proportional to $\cos \phi_0$ and $\sin \phi_0$, respectively, where ϕ_0 is the true value of ϕ of a target whose position is being estimated. The second part of the criterion is that the standard deviations of X and Y be made small in comparison to the means, as the means assume their desired values.

In examining the quality of the decision, only X will be analyzed, but the analysis of Y is essentially identical. The choice of radar parameters to optimize X will also optimize Y.

Minimization of Interference

Consider a target at (Δ_0, ϕ_0) with respect to the scan axis which returns a signal-to-noise ratio $a(\Delta_0, \phi_0, \psi_n)$. Let

$$z_n = \frac{y_n^2}{2\sigma^2[1 + a(\Delta_0, \phi_0, \psi_n)]}. \tag{15}$$

Then $z_n = \frac{1}{2}\kappa_2$ where κ_2 has a chi-square distribution of two degrees of freedom. The mean of z_n is then one, as is also the variance. Using the approximation of (10), X may be written

$$X = \sum_{n=1}^{n=k} 2\sigma^2 z_n\left[1 + a_0 + \frac{a_0 2.8\theta_0\Delta_0}{B^2}\cos(\psi_n - \phi_0)\right]\cos\psi_n$$

$$= 2\sigma^2(1 + a_0)\sum_{n=1}^{n=k} z_n(\cos\psi_n + \cos\psi_n \sin\psi_n b \sin\phi_0 + \cos^2\psi_n b\cos\phi_0), \quad (16)$$

where

$$b = \frac{a_0}{1 + a_0}\,\frac{2.8\theta_0\Delta_0}{B^2}. \quad (17)$$

Since the z_n's are independent, the mean value of X is

$$\overline{X} = 2\sigma^2(1 + a_0)\sum_{n=1}^{n=k}(\cos\psi_n + \cos\psi_n\sin\psi_n b\sin\phi_0 + \cos^2\psi_n b\cos\phi_0). \quad (18)$$

In accordance with the chosen criterion of quality, the first two terms following the summation sign are interference and should be minimized. The first term is made zero if the pulses are oppositely paired. That is, for every pulse transmitted at ψ_n, one is also transmitted at $\psi_n + \pi$. The second term, which may be considered cross-talk, is made zero if pulses are arranged in quadruple sets. Then, for each pulse at ψ_n, pulses are also transmitted at $\psi_n + \pi/2$, $\psi_n + \pi$, and $\psi_n + 3\pi/2$. With this pulse pattern, $\overline{X}$ becomes

$$\overline{X} = 2\sigma^2(1 + a_0)\sum_{n=1}^{n=k/2}(\cos^2\psi_n + \sin^2\psi_n)b\cos\phi_0$$

$$= 2\sigma^2(1 + a_0)\,k/2\; b\cos\phi_0. \quad (19)$$

Now consider the variance of X.

$$\sigma_X^2 = [2\sigma^2(1 + a_0)]^2\sum_{n=1}^{n=k}(\cos\psi_n + \cos\psi_n\sin\psi_n b\sin\phi_0 + \cos^2\psi_n b\cos\phi_0)^2$$

$$= [2\sigma^2(1 + a_0)]^2\sum_{n=1}^{n=k}(\cos^2\psi_n + \cos^2\psi_n\sin^2\psi_n b^2\sin^2\phi_0 + \cos^4\psi_n b^2\cos^2\phi_0 + 2\cos^3\psi_n b\cos\phi_0 + 2\cos^2\psi_n\sin\psi_n b\sin\phi_0 + 2\cos^3\psi_n\sin\psi_n b^2\sin\phi_0\cos\phi_0). \quad (20)$$

For the quadruple-set pattern of pulsing, the last three terms of σ_X^2 become zero. The first term becomes $k/2$. The second may be written

$$\sum_{n=1}^{n=k}\cos^2\psi_n\sin^2\psi_n b^2\sin^2\phi_0 = 4\sum_{n=1}^{n=k/4}\cos^2\psi_n\sin^2\psi_n b^2\sin^2\phi_0, \quad (21)$$

where, for n from 1 to $k/4$, the ψ_n's are in one quadrant.

The third term may be written

$$\sum_{n=1}^{n=k}\cos^4\psi_n b^2\cos^2\phi_0 = \sum_{n=1}^{n=k/2}(\cos^4\psi_n + \sin^4\psi_n)b^2\cos^2\phi_0$$

$$= k/2\; b^2\cos^2\phi_0 - 4\sum_{n=1}^{n=k/4}\cos^2\psi_n\sin^2\psi_n b^2\cos^2\phi_0, \quad (22)$$

where the n's are chosen from the appropriate quadrants.

Combining terms,

$$\sigma_X^2 = [2\sigma^2(1 - a_0)]^2[k/2(1 + b^2\cos^2\phi_0) - 4b^2(\sin^2\phi_0 - \cos^2\phi_0)\sum_{n=1}^{n=k/4}\cos^2\psi_n\sin^2\psi_n. \quad (23)$$

The last term in (23) can be positive, negative, or zero, depending upon ϕ_0. However, the corresponding term which arises in the analysis of Y has the opposite sign. Therefore, the term ought to be eliminated. It will be, if pulses are transmitted only at 0, $\pi/2$, π, and $3\pi/2$. These four values of ψ_n are optimum for transmission, if such an arrangement does not limit another parameter. In a system where the time consumed in switching the beam axis from one value of ψ_n to another is not negligible, the number of pulses k available for an observation are limited. In such a system, the necessity for increasing k makes it necessary to choose adjacent ψ_n's close together. In any event, because of our prior assumption of very small Δ/B, b would be small enough to make the latter term in the brackets very small in comparison to the first term.

Assuming, however, that pulses are transmitted only at 0, $\pi/2$, π, and $3\pi/2$, the ratio of standard deviation to the mean becomes

$$\frac{\sigma_X}{\overline{X}} = \frac{\sqrt{2}\sqrt{1 + b^2\cos^2\phi_0}}{\sqrt{k}\; b\cos\phi_0}. \quad (24)$$

This quantity is reduced with increasing k and increasing b. The assumption that was made in approximating $a(\Delta, \phi_0, \psi_n)$ also makes b very small in comparison to one. Hence, how large we can make b and still improve performance depends upon an analysis with a better approximation of $a(\Delta, \phi, \psi_n)$. The method of decision in (14) could be tested for validity in the case of large b by using the exact expression of $a(\Delta, \phi, \psi_n)$ in (16) and the subsequent analysis. A useful series for $a(\Delta_0, \phi_0, \psi_n)$ in that case would be

$$a(\Delta_0, \phi_0, \psi_n) = a_0\exp\left[-\frac{1.4\Delta_0^2}{B^2}\right]\cdot\left[I_0(r) + 2\sum_{p=1}^{p=\infty} I_p(r)\cos p(\psi_n - \phi_0)\right], \quad (25)$$

where $I_\nu(r)$ is the modified Bessel function of the first kind and

$$r = \frac{2.8\Delta_0}{B^2}. \quad (26)$$

However, such a test will not be made here.

At any rate, if b is very small, there is a tendency toward improvement if b is increased. This can be done for constant Δ_0 by increasing θ_0 and decreasing B, as long as a_0 (which is a function of θ_0 and B) is kept large in comparison to one.

Reduction of Interference by Pulse Correlation

An additional method of reducing interference is suggested, in the case of a $0 - \pi/2 - \pi - 3\pi/2$ pulse pattern, by writing X in this fashion:

$$X = 2\sigma^2(1 + a_0) \sum_{i=1}^{i=k/4} z_i(b \cos \phi_0 + 1) + \sum_{j=1}^{j=k/4} z_j(b \cos \phi_0 - 1), \qquad (27)$$

where i is for times when $\cos \psi_n = 1$ and j is for times when $\cos \psi_n = -1$. If pulse pairs could be completely correlated so that $z_i = z_j$, the mean of X would become

$$\overline{X} = 2\sigma^2(1 + a_0)\, k/2\, b \cos \phi_0. \qquad (28)$$

This is the same as (19). The variance is now

$$\sigma_X^2 = [2\sigma^2(1 + a_0)]^2\, k\, b^2 \cos^2 \phi_0, \qquad (29)$$

and the ratio of the standard deviation to the mean

$$\frac{\sigma_X}{\overline{X}} = \frac{2}{\sqrt{k}}. \qquad (30)$$

For small b, this is an improvement over (24).

It is possible that close correlation could be established by simultaneous, or nearly simultaneous, transmission of pulses at opposite sides of the beam locus. Even if the received signals were completely correlated in amplitude for opposing-pulse pairs, however, the added noise would still be uncorrelated. Therefore, $\sigma_x/\overline{X}$ would be larger than (30), depending upon the magnitude of a_0 and $b \cos \phi_0$, but smaller than (24).

Conclusion

The conclusions to be drawn from the above discussion are: 1) Interference due to noise and signal fluctuation is reduced if pulses are paired to be transmitted on opposite sides of the beam-axis transmission locus. 2) Interference due to cross-talk is reduced if pulses are transmitted in quadruple sets and will be zero if pulses are transmitted only at $\psi_n = 0, \pi/2, \pi$, and $3\pi/2$. 3) Interference independent of the position of the target is reduced if the pulse-pairs are transmitted simultaneously or nearly simultaneously. This might be accomplished with a rapid lobe-switching scheme. 4) The accuracy of position determination increases with the number of pulses integrated. Because of relative motion between the scan axis and the target, several radars might have to be used together to increase the effective number of pulses.

Correspondence

The Linear, Input-Controlled, Variable-Pass Network

The writer, finding himself unable to establish the validity of Theorem I-A of this paper[1] along the lines of the proof outlined there (integration by parts and use of the mean value theorem for integrals), wishes to submit an elementary proof of this theorem under more general conditions.

Original Statement of Theorem I-A

If $F(t)$ is a bounded, continuous, single-valued function of t, and if its derivative, $F'(t)$, is continuous, then

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} F(t)\, dt = \lim_{T\to\infty} \frac{\alpha}{T} \int_{-T}^{T} \int_{-\infty}^{t} F(x) e^{-2\alpha(t-x)}\, dx\, dt$$

$$0 < \alpha < \infty, \qquad \alpha \text{ real}.$$

[1] B. E. Keiser, "The linear input-controlled, variable-pass network," Trans. IRE, vol. IT-1, pp. 34–39; March, 1955.

The result of this communication may be stated as follows.

New Statement of Theorem I-A

If $F(t)$ is a bounded, continuous, single-valued function of t, then

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} F(t)\, dt = \lim_{T\to\infty} \frac{\alpha}{T} \int_{-T}^{T} \int_{-\infty}^{t} F(x) e^{-2\alpha(t-x)}\, dx\, dt$$

$$0 < \alpha < \infty, \qquad \alpha \text{ real}.$$

Proof: Consider function $I(T)$ defined by

$$I(T) = \frac{\alpha}{T} \int_{-T}^{T} \int_{-\infty}^{t} F(x) e^{-2\alpha(t-x)}\, dx\, dt. \qquad (1)$$

The domain of integration D of $I(T)$ is shown in the accompanying figure where $D = D_1 + D_2$ is a convenient partitioning of D into subdomains as shown. By consideration of (1) and the partitioning of D we may at once write

$$I(T) = I_1(T) + I_2(T), \qquad (2)$$

where

$$I_1(T) = \frac{\alpha}{T} \iint_{D_1} F(x) e^{-2\alpha(t-x)}\, dx\, dt \qquad (3)$$

and

$$I_2(T) = \frac{\alpha}{T} \iint_{D_2} F(x) e^{-2\alpha(t-x)}\, dx\, dt. \qquad (4)$$

Consider first the function $I_1(T)$. An examination of Fig. 1 shows that we may write

$$I_1(T) = \frac{\alpha}{T} \int_{-T}^{T} \int_{-T}^{t} F(x) e^{-2\alpha(t-x)}\, dx\, dt. \qquad (5)$$

Interchanging the order of integration this becomes

$$I_1(T) = \frac{\alpha}{T} \int_{-T}^{T} \int_{x}^{T} F(x) e^{-2\alpha(t-x)}\, dt\, dx. \qquad (6)$$

It is now possible to perform the integration with respect to t to obtain the result

$$I_1(T) = \frac{1}{2T}\int_{-T}^{T} F(x)e^{2\alpha x} \cdot (e^{-2\alpha x} - e^{-2\alpha T})\, dx, \quad (7)$$

whence

$$I_1(T) = \frac{1}{2T}\int_{-T}^{T} F(x)\, dx - \frac{e^{-2\alpha T}}{2T}\int_{-T}^{T} F(x)e^{2\alpha x}\, dx. \quad (8)$$

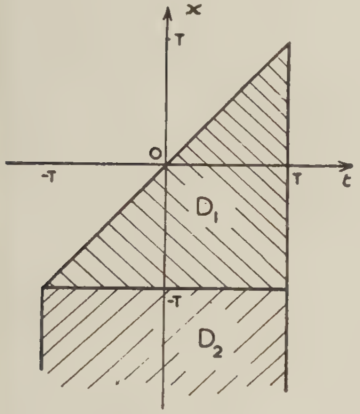


Fig. 1—Domains of integration.

Since by hypothesis $F(x)$ is bounded, with upper and lower bounds M and m respectively, say, and α is real and positive, the last term of (8), which will be denoted by $J(T)$, becomes

$$J(T) \leq \frac{Me^{-2\alpha T}}{2T}\int_{-T}^{T} e^{2\alpha x}\, dx \quad (9)$$

and

$$J(T) \geq \frac{me^{-2\alpha T}}{2T}\int_{-T}^{T} e^{2\alpha x}\, dx, \quad (10)$$

giving the result

$$\frac{m}{4\alpha T}(1 - e^{-4\alpha T}) \leq J(T) \leq \frac{M}{4\alpha T}(1 - e^{-4\alpha T}). \quad (11)$$

It follows directly from (11) that

$$\lim_{T\to\infty} J(T) = 0. \quad (12)$$

Consider next the function $I_2(T)$, which from the definition of D_2 may be written as

$$I_2(T) = \frac{\alpha}{T}\int_{-T}^{T}\int_{-\infty}^{-T} F(x)e^{-2\alpha(t-x)}\, dx\, dt. \quad (13)$$

This may at once be expressed as the product of two simple integrals to give

$$I_2(T) = \frac{\alpha}{T}\int_{-T}^{T} e^{-2\alpha t}\, dt \cdot \int_{-\infty}^{-T} F(x)e^{2\alpha x}\, dx, \quad (14)$$

whence

$$I_2(T) = \frac{e^{2\alpha T}}{2T}[1 - e^{-4\alpha T}] \cdot \int_{-\infty}^{-T} F(x)e^{2\alpha x}\, dx. \quad (15)$$

Employing an argument similar to that used when considering $J(T)$, we find at once that

$$\frac{m}{4\alpha T}[1 - e^{-4\alpha T}] \leq I_2(T) \leq \frac{M}{4\alpha T}[1 - e^{-4\alpha T}]. \quad (16)$$

It again follows directly that

$$\lim_{T\to\infty} I_2(T) = 0. \quad (17)$$

From (1), (2), (8), (12) and (17) it follows finally that

$$\lim_{T\to\infty} I(T) = \lim_{T\to\infty} \frac{1}{2T}\int_{-}^{T} F(x)\, dx, \quad (18)$$

which is the required result.

It may be noted here in passing that this new statement of Keiser's Theorem I-A is still far from the most general formulation. It has been quoted here requiring continuity of $F(t)$, since the above simple proof then follows directly, but provided $F(t)$ is Lebesgue integrable an appeal to Fubini's Theorem establishes the result under far more general conditions on $F(t)$.

ALAN JEFFREY
G.E.C. Stanmore Labs.
Middlesex, England

Rebuttal

The proof offered by Mr. Jeffrey is correct in every respect, and I raise no question about the validity of his comments. However, the functions which the above-referenced paper deals with have derivatives which are bounded and continuous everywhere. Therefore, no attempt was made to obtain a more general proof.

The proof of the theorem was omitted because of lack of space. If either the First or the Second Mean Value Theorem is utilized to obtain the proof, then, of course, $F'(t)$ must be bounded everywhere, as well as continuous.

Mr. Jeffrey has submitted to me a proof using the First Mean Value Theorem. My proof, which uses the Second Mean Value Theorem, also requires that $F'(t)$ be bounded everywhere and continuous, although this restriction can be lifted with a proof of the type presented by Mr. Jeffrey.

The theorem and its proof are as follows:

Theorem I-A

If $F(t)$ is a bounded, continuous, single-valued function of t, and if $F'(t)$ is continuous and bounded, then

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} F(t)\, dt = \lim_{T\to\infty}\frac{\alpha}{T}\int_{-T}^{T}\int_{-\infty}^{t} F(x)e^{-2\alpha(t-x)}\, dx\, dt.$$

Proof: Consider the quantity

$$2\alpha\int_{-\infty}^{t} e^{-2\alpha t}e^{2\alpha x}F(x)\, dx.$$

Let $F(x) = u$ and let $2\alpha e^{2\alpha x}\, dx = dv$. Since

$$\int u dv = uv - \int v du,$$

$$e^{-2\alpha t}\int_{-\infty}^{t} 2\alpha e^{2\alpha x}F(x)\, dx = e^{-2\alpha t}\left[F(x)e^{2\alpha x}\Big|_{-\infty}^{t} - \int_{-\infty}^{t} e^{2\alpha x}F'(x)\, dx\right] = e^{-2\alpha t}\left[F(t)e^{2\alpha t} - \int_{-\infty}^{t} e^{2\alpha x}F'(x)\, dx\right].$$

Thus

$$e^{-2\alpha t}\int_{-\infty}^{t} 2\alpha e^{2\alpha x}F(x)\, dx = F(t) - \int_{-\infty}^{t} e^{-2\alpha(t-x)}F'(x)\, dx. \quad (1)$$

Integrate each side with respect to t from $t = 0$ to $t = T$, divide by T, and take the limit of each side as $T \to \infty$. This gives

$$\lim_{T\to\infty}\frac{1}{T}\int_{0}^{T} e^{-2\alpha t}\int_{-\infty}^{t} 2\alpha e^{2\alpha x}F(x)\, dx\, dt = \lim_{T\to\infty}\frac{1}{T}\int_{0}^{T} F(t)\, dt - \lim_{T\to\infty}\frac{1}{T}\int_{0}^{T}\int_{-\infty}^{t} e^{-2\alpha(t-x)}F'(x)\, dx\, dt. \quad (2)$$

The following is a proof that, under the conditions postulated,

$$\lim_{T\to\infty}\frac{1}{T}\int_{0}^{T}\int_{-\infty}^{t} e^{-2\alpha(t-x)}F'(x)\, dx\, dt = 0. \quad (3)$$

First note that

$$\lim_{T\to\infty}\frac{1}{T}\int_{0}^{T}\int_{-\infty}^{t} e^{-2\alpha(t-x)}F'(x)\, dx\, dt = \lim_{T\to\infty}\frac{1}{T}\int_{0}^{T} e^{-2\alpha t} \cdot \left[\int_{-\infty}^{t} e^{2\alpha x}F'(x)\, dx\right] dt. \quad (4)$$

Consider the term in brackets on the right-hand side,

$$\int_{-\infty}^{t} e^{2\alpha x} F'(x)\, dx.$$

There is a mean-value theorem for integrals[2] which states:

Let $f(x)$ and $\varphi(x)$ be two functions which are continuous in (a, b). If $\varphi(x)$ is a positive, monotonic function, increasing in (a, b), then there exists a value ξ, where $a \leq \xi \leq b$, such that

$$\int_{a}^{b} f(x)\varphi(x)\, dx = \varphi(b) \int_{\xi}^{b} f(x)\, dx. \quad (5)$$

Let $\varphi(x) = e^{2\alpha x}$ and let $f(x) = F'(x)$. Take a as $-\infty$ and b as t. Then the theorem quoted above states that there is a ξ between $-\infty$ and t such that

$$\int_{-\infty}^{t} e^{2\alpha x} F'(x)\, dx = e^{2\alpha t} \int_{\xi}^{t} F'(x)\, dx. \quad (6)$$

Integration of the right-hand side yields

$$\int_{-\infty}^{t} e^{2\alpha x} F'(x)\, dx = e^{2\alpha t}[F(t) - F(\xi)]. \quad (7)$$

Substituting this relationship into the right-hand side of (4) gives

$$\lim_{T\to\infty} \frac{1}{T} \int_{0}^{T} \int_{-\infty}^{t} e^{-2\alpha(t-x)} F'(x)\, dx\, dt$$

$$= \lim_{T\to\infty} \frac{1}{T} \int_{0}^{T} [F(t) - F(\xi)]\, dt. \quad (8)$$

If $F(t)$ is a constant for all values of t, then $F(\xi)$ must be this same constant, since $F(\xi)$ merely takes on values which $F(t)$ took on previously. In this case, the right-hand side of (8) is zero. Suppose $F(t)$ is not a constant, but is the sum of a constant term and a term varying with t. Then $F(\xi)$ is of the same nature. In particular,

$$F(t) = A + B(t) \quad (9)$$

$$F(\xi) = A + B(\xi), \quad (10)$$

where the average value of $B(t)$ over all t, or of $B(\xi)$ over all ξ or t is zero, from the way in which the constant A was chosen. Then

$$\lim_{T\to\infty} \frac{1}{T} \int_{0}^{T} [F(t) - F(\xi)]\, dt \quad (11)$$

$$= \lim_{T\to\infty} \frac{1}{T} \int_{0}^{T} [B(t) - B(\xi)]\, dt = 0.$$

Therefore, from (8),

$$\lim_{T\to\infty} \frac{1}{T} \int_{0}^{T} \int_{-\infty}^{t} e^{-2\alpha(t-x)} F'(x)\, dx\, dt = 0. \quad (12)$$

From (2), it follows that

$$\lim_{T\to\infty} \frac{1}{T} \int_{0}^{T} F(t)\, dt$$

$$= \lim_{T\to\infty} \frac{2\alpha}{T} \int_{0}^{T} \int_{-\infty}^{t} e^{-2\alpha(t-x)} F(x)\, dx\, dt. \quad (13)$$

Since $F(t)$ is assumed to be a bounded, continuous function of t for all values of t, the integral on the left-hand side of (13) also exists if its limits are taken as $-T$ to T. Furthermore, $\int_{-\infty}^{t} e^{-2\alpha(t-x)} F(x)\, dx$ is finite for all values of t. Therefore, in the integration with respect to t on the right-hand side of (13), the limits 0 to T may be replaced by $-T$ to T, and the resulting integral still will converge. Thus, in view of the fact that $F(t)$ is bounded for negative values of t as well as for positive values of t, it is permissible to write

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} F(t)\, dt$$

$$= \lim_{T\to\infty} \frac{\alpha}{T} \int_{-T}^{T} \int_{-\infty}^{t} e^{-2\alpha(t-x)} F(x)\, dx\, dt. \quad (14)$$

B. E. Keiser
White-Rogers Electric Co.
St. Louis 6, Mo.

[2] I. S. Sokolnikoff, "Advanced Calculus," McGraw-Hill Book Co., Inc., New York, N. Y., p. 115; 1949.

Contributors

Nelson M. Blachman (A'50—SM'53) was born in Cleveland, Ohio, on October 27, 1923. He received his B.S. in physics in 1943 from the Case Institute of Technology, and his A.M. in physics and Ph.D. in engineering sciences and applied physics in 1947 from Harvard University, where he was a John Tyndall and a Gordon McKay scholar in the respective departments.


N. M. Blachman

From 1943 to 1945 he was a member of the Theory and Transducer Groups at the Harvard Underwater Sound Laboratory, where he was concerned with the design and measurement of electroacoustic transducers and with the analysis of sonar system designs. From 1945 to 1946, Dr. Blachman worked at the Cruft Laboratory, Harvard University, on signal and noise problems in radio communication, particularly fm.

As a member of the Theory Group of the Accelerator Project at the Brookhaven National Laboratory from 1947 to 1951, Dr. Blachman was concerned with the theory and design of the Cosmotron, Brookhaven's three-billion-volt proton synchrotron. From 1951 to 1954, he was a member of the staff of the Mathematical Sciences Division of the Office of Naval Research, Washington, D. C., administering ONR's program of supported research in the fields of computers and mathematics. In this position, he compiled a number of surveys of various aspects of the digital-computer field, including the 1953 ONR survey of automatic digital computers.

In 1954, Dr. Blachman joined the Systems Branch of Sylvania's Electronic Defense Laboratory, Mountain View, Calif., where he has been concerned mainly with communication theory and the design of electronic countermeasure systems.

Dr. Blachman is the author of numerous scientific papers, principally in the fields of communication theory and synchrotron theory, as well as the editor of several documents on computers published by the Office of Naval Research. He is a member of the American Physical Society, the Association for Computing Machinery, Sigma Xi, the Federation of American Scientists, the Institute of Mathematical Statistics, and the Scientific Research Society of America.

❖

Gerald P. Dinneen was born in Elmhurst, N. Y. on October 23, 1924. He received the B.S. degree in mathematics from Queens College in 1947, the M.S. and Ph.D. degrees in mathematics from the University of Wisconsin in 1948 and 1952.


G. P. Dinneen

From 1943 to 1946, Dr. Dinneen served with the U. S. Air Force in the South Pacific theatre. After leaving Wisconsin, he worked for the Aerophysics department of Goodyear Aircraft. Since January, 1953,

he has been a staff member of the Lincoln Laboratory of the Massachusetts Institute of Technology.

Dr. Dinneen is a member of Signa Xi and the American Mathematical Society.

❖

Behrend J. DuWaldt was born on September 26, 1927 in Chicago, Ill. He received the B.S. degree from the U. S. Naval Academy in 1949. After three years of active duty as a line officer with the Navy, he enrolled in the graduate school of Purdue University to major in communications engineering. He was awarded the M.S.E. degree in 1954 and the Ph.D. degree in 1955.

B. J. DuWalt

He engaged in transistor circuitry research at Purdue as a graduate assistant in 1953 and 1954, and was appointed research instructor from 1954 to 1955. He joined the Communications Division of the Ramo-Wooldridge Corp. in 1955.

Dr. DuWaldt is a member of Eta Kappa Nu and Sigma Xi.

❖

Douglas G. Lampard (M '55) was born in Sydney, Australia on May 4, 1927. He received the B.S. degree with first class honors in physics in 1951, and the M.S. degree in 1952, both from the University of Sydney. From 1951 to 1952, he was a member of the microwave group of the C.S.I.R.O. Division of Electrotechnology, Sydney, where he did developmental work in microwave spectroscopy and also worked on noise theory.

D. G. Lampard

In 1952, he obtained a C.S.I.R.O. overseas studentship which enabled him to continue his work on noise theory at the University of Cambridge, England. He obtained the Ph.D. degree from the University of Cambridge in 1954 and from September, 1954 to January, 1955 was visiting lecturer in the department of electrical engineering, Columbia University, New York. Since then Dr. Lampard has returned to the C.S.I.R.O. Division of Electrotechnology where he is working on electrostatics, noise theory and related problems.

Dr. Lampard is an associate of the Cambridge Philosophical Society and of the Institute of Physics.

❖

Alan B. Lees was born on October 24, 1926, in Manchester, England. He received his B.Sc.(Tech.) and M.Sc.(Tech.) degrees in Electrical Engineering in 1948 and 1949, and his Ph.D. in 1954, all from the University of Manchester.

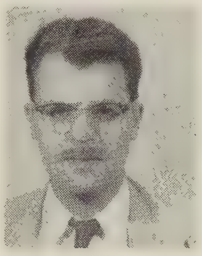
A. B. Lees

In 1949 and 1950, he was a development engineer with Ferranti, Ltd. From 1950 to 1954, Dr. Lees was a Lecturer in Electrical Engineering at the University of Manchester. In 1954-55, he was Visiting Lecturer in Electrical Engineering at the Massachusetts Institute of Technology; and in 1955-56 has been Visiting Lecturer in E.E. at Columbia University, all while on leave from the University of Manchester as a visiting Fulbright lecturer.

Dr. Lees is a member of the Institution of Electrical Engineers.

❖

Irving S. Reed was born in Seattle, Wash. on November 12, 1923. He attended the University of Alaska for two years and graduated from the California Institute of Technology in 1944 with the B.S. degree. He received the Ph.D. degree in mathematics from the California Institute of Technology in 1949.



I. S. Reed

From 1944 to 1946 Dr. Reed was in the U. S. Navy. From 1947 to 1950 he worked for Northrop Aircraft, Inc., in Hawthorne, Calif., during which he contributed to the development of the Maddida computer. In 1950 Dr. Reed became one of the founding directors of the Computer Research Corp., and in October, 1951 he joined the Lincoln Laboratory of the Massachusetts Institute of Technology.

Dr. Reed is a member of the American Mathematical Society.

❖

Herbert Sherman (S'40—A'41—SM'49) was born on February 24, 1920 in New York, N. Y. He received the bachelor's degree in electrical engineering from the College of the City of New York, in 1940. He received the degrees of Master of Electrical Engineering in 1949 and Doctor of Electrical Engineering in June, 1955 from the Polytechnic Institute of Brooklyn.

Herbert Sherman

After receiving his undergraduate degree he was employed by the Army Signal Corps in the inspection and production planning of electronic equipment for military use. In 1944 he left the Signal Corps, and was commissioned in the U. S. Naval Reserve. His principal military assignment was with the Operational Development Force as a radar officer aboard the U.S.S. Portsmouth.

In 1946, Dr. Sherman joined the Air Force Watson Laboratories at Red Bank, N. J. (later the Rome Air Development Center) as a member of the Plans Staff. There he was engaged in systems planning for major ground-based Air Force electronic systems, and in establishing the development program for their implementation. Later he became chief of the Plans Section of the Laboratories, and also served on various panels and committees of the Research and Development Board as an Air Force representative.

In 1952 he became a staff member of the Lincoln Laboratory, M.I.T., where he is currently leader of a group developing communications for the SAGE System. He also serves on the panel of consultants of the Technical Advisory Panel on Electronics to the Assistant Secretary of Defense for Research and Development.

Dr. Sherman is a member of the American Institute of Electrical Engineers, the Association for Computing Machinery, the American Association for the Advancement of Science, Sigma Xi, and Tau Beta Pi.

INFORMATION FOR AUTHORS

Authors are requested to submit editorial correspondence or technical manuscripts to the Publications Chairman for possible publication in the PGIT TRANSACTIONS. Papers submitted should include a statement as to whether the material has been copyrighted, previously published, or accepted for publication elsewhere.

Papers should be written concisely, keeping to a minimum all introductory and historical material. It is seldom necessary to reproduce in their entirety previously published derivations, where a statement of results, with adequate references, will suffice.

To expedite reviewing procedures, it is requested that authors submit the original and two legible copies of all written and illustrative material. The manuscript should be double-spaced, and the illustrations drawn in india ink on drawing paper or drafting cloth. Each paper should include a carefully written abstract of not more than 200 words. Upon acceptance, papers should be prepared for publication in a manner similar to those intended for the PROCEEDINGS OF THE IRE. Further instructions may be obtained from the Publications Chairman. Material not accepted for publication will be returned.

IRE TRANSACTIONS ON INFORMATION THEORY is published three times a year, in March, September, and December. A minimum of one month must be allowed for review and correction of all accepted manuscripts. A period of approximately two months additional is required for the mechanical phases of publication and printing. Therefore, all manuscripts must be submitted three months prior to the respective publication dates. There is no June issue. The IRE CONVENTION RECORD is published at that time, and a bound collection of Information Theory papers delivered at the annual IRE National Convention is mailed gratis to all PGIT members.

All technical manuscripts and editorial correspondence should be addressed to Laurin G. Fischer, Federal Telecommunications Labs., 492 River Road, Nutley, N. J. Local Chapter activities and announcements, as well as other nontechnical news items, should be addressed to Nathan Marchand, Marchand Electronic Labs., Riversville Road, Greenwich, Conn.